

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2004, 21:22:32 ; Search time 5635 Seconds

(without alignments)  
4684.278 Million cell updates/sec

Title: US-09-890-475-1

Perfect score: 3104  
Sequence: 1 MSNVPPTVAAQPTTANPLL.....RYLGSLNQRSPRNSSLDPX 609

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgm2\_1/USPRO\_spool\_P/US09890475/runac\_24022004\_135311\_5291/app\_query.fasta\_1.775  
-DB=genbml -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=ptco -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09890475 @CGN\_1\_14958 @runac\_24022004\_135311\_5291 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genbml:\*  
1: gb\_ba:\*  
2: gb\_ncg:\*  
3: gb\_in:\*  
4: gb\_on:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sbs:\*  
28: em\_un:\*

29: em\_vt:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3104	100.0	1830	8	AF228500	Arabidops
2	3104	100.0	2257	6	AX032763	Sequence
3	2981	96.0	3761	6	AX032762	Sequence
4	2981	96.0	3761	8	AF228499	Arabidops
5	2975	95.8	3150	8	AY198403	Arabidops
6	2975	95.8	3261	8	AY198402	Arabidops
7	2975	95.8	3739	8	AY198401	Arabidops
8	2969	95.7	3301	8	AY198405	Arabidops
9	2960	95.0	3758	8	AY198404	Arabidops
10	2936	94.6	91849	8	FM23	Arabidops
11	2936	94.6	197975	8	ATCHRIV2	Arabidops
12	931	30.0	558	8	AY092681	Arabidops
13	931	30.0	558	8	AY092683	Arabidops
14	931	30.0	558	8	AY092691	Arabidops
15	931	30.0	558	8	AY092693	Arabidops
16	927	29.9	558	8	AY092677	Arabidops
17	927	29.9	558	8	AY092679	Arabidops
18	927	29.9	558	8	AY092694	Arabidops
19	927	29.9	558	8	AY092695	Arabidops
20	927	29.9	558	8	AY092696	Arabidops
21	921	29.7	558	8	AY092678	Arabidops
22	921	29.7	558	8	AY092680	Arabidops
23	921	29.7	558	8	AY092682	Arabidops
24	921	29.7	558	8	AY092684	Arabidops
25	921	29.7	558	8	AY092685	Arabidops
26	921	29.7	558	8	AY092686	Arabidops
27	921	29.7	558	8	AY092687	Arabidops
28	921	29.7	558	8	AY092688	Arabidops
29	921	29.7	558	8	AY092689	Arabidops
30	921	29.7	558	8	AY092690	Arabidops
31	921	29.7	558	8	AY092692	Arabidops
32	845	27.2	511	8	AY176673	Brassica
33	658	21.2	503	8	AY092437	Arabidops
34	658	21.2	503	8	AY092438	Arabidops
35	658	21.2	503	8	AY092439	Arabidops
36	658	21.2	503	8	AY092440	Arabidops
37	658	21.2	503	8	AY092441	Arabidops
38	658	21.2	503	8	AY092442	Arabidops
39	658	21.2	503	8	AY092443	Arabidops
40	658	21.2	503	8	AY092444	Arabidops
41	658	21.2	503	8	AY092445	Arabidops
42	658	21.2	503	8	AY092446	Arabidops
43	658	21.2	503	8	AY092447	Arabidops
44	658	21.2	503	8	AY092448	Arabidops
45	658	21.2	503	8	AY092449	Arabidops

RESULT 1

ALIGNMENTS

AF228500  
LOCUS AF228500 1830 bp mRNA linear PLN 14-OCT-2000  
DEFINITION Arabidopsis thaliana FRIGIDA (FRI) mRNA, complete cds.  
ACCESSION AF228500  
VERSION AF228500.1 GI:10801175  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 (bases 1 to 1830)  
Johnston, U., West, J., Lister, C., Michaels, S., Amadio, R. and  
Dean, C.  
TITLE Molecular analysis of FRIGIDA, a major determinant of natural  
variation in Arabidopsis flowering time  
JOURNAL Science 290 (5490), 344-347 (2000)  
MEDLINE 20485641  
PubMed 11030654  
REFERENCE  
2 (bases 1 to 1830)  
Johnston, U. and Lister, C.  
TITLE Direct Submission  
Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre,  
Colney Lane, Norwich, Norfolk NR4 7UH, U.K.  
LOCATION/Qualifiers  
1. 1830  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultivar="ecotype H51"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/map="between mi51 and mi122"  
1. 1830  
/gene="FRI"  
1. 1830  
/gene="FRI"  
/note="late flowering gene"  
/codon\_start=1  
/product="FRIGIDA"  
/protein\_id="AA023415.1"  
/db\_xref="GI:10801175"  
/translation="MSNYPRPTVAQPTTANPLORHSEPRRELPKIVTESTSMD  
ITIGSKQOPILKSIDELAASVAVETKRPDDIQRKESITENKIDSKESNVILA  
ANNKHQPMSPPRNNVSVETVTSVSPSQTIVETSNKRGSGMCKESKELKRT  
IYANISDAKIMEELPSALKAKEPAKFLVDTIGFYIQGRAFTKESPMCSAQVSL  
ILIBSFLMPDRKGKVKIISWIKDABATAVAMRKIVMTSGIAYESIKRGMEIAL  
VACFVPSNFRSTDLIDLIRMSGNRIAGALRSQIVPMYSGIAYESIKRGMEIAL  
EAVYTFGEMDKSALVLTSLFKMSKESPERAKRKAOSLPFKKATKQLAVLSVNG  
CMETHKLDPAKELPGMOIKBOIVLEKDTLIDKEMERKASLSIMEBALAKRMVNG  
QIKRPRLSMPENPTSSSYSPITPRDRSPSGRDDDOEIKALVSYIGPSTPHRS  
KSPFNPVPLPHGIGKSVYAEHLAPNYSFGHGRHGRQYSPSLVHGQRHPLQYSP  
PIHGQOQLPYGIQVRHSPSEERYLGLSNKSPRSNSLDPK"

ORIGIN  
Alignment Scores:  
Pred. No.: 2, 77e-195 Length: 1830  
Score: 3104.00 Matches: 609  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-09-890-475-1 (1-609) x AF228500 (1-1830)

QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu 20  
DB 1 ATGTCAATTAATCCACGACGGTGGGGGCGCAACCAACGACGCGCAATTCATCTGTG 60  
QY 21 GlnATGhIeGlnSerGlnGlnAArgAArgGlnLeuProTyrIleValGlnThrGlnSer 40  
DB 61 CAGCGCATTCATCTGTAACGACGACGAGAAATTAACGAAATTTGTGAAACAGAGCT 120  
QY 41 ThrSerMetAspLleThrIleGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60

DB 121 ACAAGTAGACATTCATCGCTCATCTAATGACACCGCTCAATTTTGAATCCATAGAC 180  
QY 61 GlnLeuAlaAlaPheSerValAlaValAlaGlnThrPheLysAArgGlnPheAspLeuGln 80  
DB 181 GAATTAAGTCCGCTTTTCAGTTGAGTGAAGAAACCTTAACCGCCATTCATCTTCAG 240  
QY 81 LysHisIleGlnSerIleGlnAsnAlaIleAspSerLysLeuGlnSerAsnGlnValVal 100  
DB 241 AACCACTTCAGTCAATCGAAACCGCAATTCATTCGAACTGAGATGAACGCGTTGTC 300  
QY 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnVal 120  
DB 301 CTGCGCGCGGGAACATTAATTTCCATCCGCCGATTTATCCGCTCCGGGAACATATAC 360  
QY 121 SerValGlnThrThrValThrValSerGlnProSerGlnGlnIleValProGlnThrSer 140  
DB 361 TCTGTAGAAACCAACGCTCACTGAGCCCAACGCTCAGAGATTGTACCGGAGAGCTGC 420  
QY 141 AsnLysProGlnGlnGlnLysArgMetCysGlnLeuMetCysSerLysGlnLysArgLysTyr 160  
DB 421 AATTAACCGAGGGGAGCGATGTGTGATGTGATGTGATGCAAGAGCTTCGTAATAC 480  
QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlnGlnIleProSerAlaLeuLys 180  
DB 481 ATATAGCGCAATATCTCTGATCAAGCTAATGTAAGAGATTCTTCAGCTTGAA 540  
QY 181 LeuAlaAlaGlnProAlaLysPheValLeuAspCysIleGlnLysPheTyrLeuGlnGly 200  
DB 541 TTGGCCAGAGGACGAGAAATTTGATTGATGATTGATGGCAAGTTTACTTACAGAG 600  
QY 201 ArgAArgAlaPheThrLysGlnSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220  
DB 601 CGTAGAGCATTTCTAAGAGTGCCTATGAGCTTCGACAGCAAGTTTCGCTTTATA 660  
QY 221 LeuGlnSerPheLeuLeuMetProAspArgGlnGlyLysGlnValLysIleGlnSerTyr 240  
DB 661 CTGAGCTTTTCTTAAATGCTGATCGTGAAGGAGAGAGTGAAGATTGAGAGATTGG 720  
QY 241 IleLysAspGlnAlaGlnThrAlaAlaValAlaAlaTyrArgLysArgLeuMetThrGlnGly 260  
DB 721 AATTAAGATGAGCGAGAGAGCGCTGCTGTGCTGAGGAAAGATTGATGACTAGAGA 780  
QY 261 GlnLeuAlaAlaAlaGlnLysMetAspAlaArgGlnLysLeuLeuLeuValAlaCysPhe 280  
DB 781 GGAATTGCTGGCGCTAGAGAAATGATGCAAGGGGTTTCTTTACTAGTTGCTGTGTTT 840  
QY 281 GlnValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlnSer 300  
DB 841 GGTGTCCTTCAAACTTTAGAGTACAGATTGCTGATTGATTAAGATGAGTGGATTGG 900  
QY 301 AsnGlnIleAlaGlnAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIle 320  
DB 901 AATGAGATTCGCGCTGCTTGAAGCGGTCAAGATTCTTGTCCCTAGCTTCAGATTA 960  
QY 321 ValGlnSerSerIleLysAArgGlnMetHisIleGlnAlaLeuGlnMetValTyrThrPhe 340  
DB 961 GTTGAAATCAAGTACAGCGTGAATGCAATTAAGCTCTTGAGATGGATTATACCTTT 1020  
QY 341 GlyMetGlnAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys 360  
DB 1021 GGCATGAGGATAAGTTTCAAGCTGCTAGTCTTAATCTTCAATTTTAAGATGACCAAG 1080  
QY 361 GlnSerPheGlnArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGlnAlaAla 380  
DB 1081 GAGTCATTGAGGGGCAAAACGAAAGCCCACTCAACCTGCAATTAAAGAGGAGCT 1140  
QY 381 ThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGlnThrHisLysLeuAsp 400  
DB 1141 ACAAGAGAGTACCTGTTATATCATCATGATTATGAGATGAGATCAACAGTTAGAT 1200  
QY 401 ProAlaAlaGlnLeuProGlnTyrPheGlnIleLysGlnGlnIleValSerLeuGlnLysAsp 420

Db 1201 CCTCGAAGAACTACACGAGTGGCAGATCAAGAGCAAAATTGTTAGCTTGGAGAAAGAC 1260  
Qy 421 ThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGlu 440  
Db 1261 ACTCTTCAGCTCGCAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTCTTAATGGAGGAA 1320  
Qy 441 AlalaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMet 460  
Db 1321 GCCCACTTGCACAGAAATGTATACCAACAGATAAAAGCTCAAGGTTGTCAACCCATG 1380  
Qy 461 GluMetProValThrSerSerTyrSerProIleTyrArgAspArgSerPhePro 480  
Db 1381 GAATGCCACCAAGTAATCTTCTCATCGTATCTCTATCTACCGTGATGAAGCTTTCCT 1440  
Qy 481 SerGlnArgAspAspGlnAspGluLysSerAlaLeuValSerSerTyrLeuGlyPro 500  
Db 1441 AGTCAAAAGAGACGATGACCAAGATGAATATACAGCTCTTGAGTAGTACCTACCTCGGCCG 1500  
Qy 501 SerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProLeuProHis 520  
Db 1501 TCACATCTTTCTCATCGCTCAAGAGATCCCGGAATATATGGTTCCACATCCACAT 1560  
Qy 521 GlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerPro 540  
Db 1561 GGTGGTTAGGAGAAAGTATATGTCATATGAACATCTGGCCCCAAATTCATCTCTCCA 1620  
Qy 541 GlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHis 560  
Db 1621 GGTCAAGAGACATAGACTTCATCGAGTACTCTCGCTCTTGGTTTCAGACAGACAT 1680  
Qy 561 ProLeuGlnTyrSerProIleHisGlyGlnGlnGlnLeuProTyrGlyIleGlnArg 580  
Db 1681 CCACATAGTACTCTCTCCAAATCATGGAACAACAGATGATACCATATGATGATACAAAG 1740  
Qy 581 ValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerPro 600  
Db 1741 GTTTACAGACATTCACCATCTGGAAGAGATTTTGGTTTATCCAAATCAAGTCTCTCT 1800  
Qy 601 ArgSerAsnSerLeuAspProLys 609  
Db 1801 CGCAGTAATCATATTAGACCCCAAA 1827

RESULT 2  
AX032763 2257 bp DNA linear PAT 21-SEP-2000  
LOCUS Sequence 3 from Patent WO0046358.  
DEFINITION AX032763  
ACCESSION AX032763  
VERSION AX032763.1 GI:10279739  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences  
REFERENCE 1  
AUTHORS Dean,C., West,J. and Johanson,U.  
TITLE plant gene  
JOURNAL Patent: WO 0046358-A 3 10-AUG-2000;  
DEAN CAROLINE (GB); WEST JOANNE (GB); PLANT BIOSCIENCE LTD (GB);  
JOHANSON URBAN (SE)  
FEATURES  
source Location/Qualifiers  
1..2257  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="likely cDNA sequence of the H51 FRI gene"

ORIGIN  
Alignment Scores:  
Pred. No.: 3,64e-195 Length: 2257  
Score: 3104.00 Matches: 609  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-890-475-1 (1-609) x AX032763 (1-2257)  
Qy 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu 20  
Db 362 ATGTCCAATTTATCCACCGAGTGGCGCGCAACCCACACGACGGCGAATCCCATGCTG 421  
Qy 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40  
Db 422 CAGCGACATCAATCTGAACAGCGACAGAGAAATACCGAAGATTGTGGAACAGAGTCT 481  
Qy 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60  
Db 482 ACAAGTATGACATTCAGATCGTCAATCTAAGCAGCCTCAATTTTGGAAATCCATAGAC 541  
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80  
Db 542 GAATTAGCTGCGCTTTTCAGTTGCGATGGAAACATTTCAACCGCAATTCGATGATCTTCAG 601  
Qy 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100  
Db 602 AAGCACATCGAGTCAATCGMAAACGAATTTGATTCGAACTCGAGAGTAACCGCGCTTGT 661  
Qy 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120  
Db 662 CTCGGCGCGCGAACAATAATTTTCATCAGCCGATGTTATCCCTCCCGCGAACATGTA 721  
Qy 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140  
Db 722 TCTGTAGAAACCAACCGCTCATCTGTAGCCAAACCGTCTCAGGAGATTGTACCGGAGACGTC 781  
Qy 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160  
Db 782 AATAAACCGGAGCGGCGATGTTGTGAGTTGATGTGTAGCAAAAGGTTCTCGGTAAATAC 841  
Qy 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180  
Db 842 ATATACCGCAATATCTCTGATCACTAAGTTAATGGAAGATTCTTCACGTTTGAAA 901  
Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200  
Db 902 TTGGCCAGAGAGCCAGCAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTACAAGG 961  
Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220  
Db 962 CGTAGAGCAATTTACTAAAGAGTCGCTATGAGCTCTGCGAGACAAGTTTCGTTCTTATA 1021  
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrp 240  
Db 1022 CTGAGTCTTTCTTCTAATGCTGATCGTGTGTAAGGGAAGGTGAAGATTGAGAGTTGG 1081  
Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGly 260  
Db 1082 ATTAAGATGAGCGGAGACGCTGCTGTGTTGGAGGAAAGGTTGAAGTGAAGGA 1141  
Qy 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280  
Db 1142 GGATTAGCTGCGCTGAGAAATGGATGCAAGGGTTTCTTTTACTAGTTGCTGTTTT 1201  
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300  
Db 1202 GGTGTTCTTCAAACTTTTAGGAGTACAGATTTGCTGATTTGATAGGATGAGTGGTTCG 1261  
Qy 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIle 320  
Db 1262 AATGAGATTGCGGCTGCTTTGAAGCGGTCAAGTTTCTTGTCCCTATGCTCAGGTATA 1321  
Qy 321 ValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPhe 340  
Db 1322 GTTGAATCAAGTATCAAGCGTGAATGTCATATTGAAGCTCTTGAAGTGGTTTATACCTTT 1381  
Qy 341 GlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys 360

```

Db      1382  GGCATGGAGGATAAGTTTTCAGCTGCTCTAGTTCTTAATCTTAAAGATGAGCAAG 1441
Qy      361  GluSerPheGluAArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAla 380
Db      1442  GAGTCATTTTGAGGGGCAAAACGGAAGCCAGTCAGCTGGCATTTAAAGAGCGGCT 1501
Qy      381  ThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAsp 400
Db      1502  ACAAGACAGCTAGCTGTGTTATCATCATGTTATGAGTGTATGAGAGCTCAAGTTAGAT 1561
Qy      401  ProAlaLysGluLeuProGlyTTPGlnLileLysGluGlnLileValSerLeuGluLysAsp 420
Db      1562  CCTGCGAAGAACATACCAAGATGGCAGATCAAGAGCAAAATTTGTTAGCTTGGAGAAAGAC 1621
Qy      421  ThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGlu 440
Db      1622  ACTCTTCAGCTCGCAAAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAAATGGAGAA 1681
Qy      441  AlaAlaLeuAlaLysArgMetTyrAsnGlnGlnLileLysArgProArgLeuSerProMet 460
Db      1682  GCGGCACTTGGCAAGAGATGTATTAACCAACAGATTAACCTCCAGGTTGTCAACCATG 1741
Qy      461  GluMetProProValThrSerSerSerTyrSerProLileTyrArgAspArgSerPhePro 480
Db      1742  GAAATGCCACAGTAATCTTTCATCGTATTTCTCTATCTACCGGTGATAGAAGCTTTCTCT 1801
Qy      481  SerGlnArgAspAspAspGlnAspGluLileSerAlaLeuValSerSerTyrLeuGlyPro 500
Db      1802  AGTCAAGAGACATGACCAAGATGAAATATCAGCTCTTGTGAGTAGTACTCTGGCCCG 1861
Qy      501  SerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuProHis 520
Db      1862  TCAACATCTTTTCTCATCGCTCAAGAGATCCCGGAATATATGGTTCCACTTCCACAT 1921
Qy      521  GlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerPro 540
Db      1922  GGTGGGTAGGAAGAGGTATATGATATGATATGATATGATATGATATGATATGATATGAT 1981
Qy      541  GlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHis 560
Db      1982  GGTCAAGGACATAGACTTCATCGACAGTACTCTCGCTCTTTGGTTTCAAGGACAGACAT 2041
Qy      561  ProLeuGlnTyrSerProProLileHisGlyGlnGlnGlnLeuProTyrGlyLileGlnArg 580
Db      2042  CCACATACAGTACTCTCTCTCCATTCATGGACAAACAGATGATACCATATGGTATACAAAG 2101
Qy      581  ValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerPro 600
Db      2102  GTTACAGACATTCACCATCTGAGAGAGATATTTGGGTTTATCCAAATCAAGGTCTCTCT 2161
Qy      601  ArgSerAsnSerSerLeuAspProLys 609
Db      2162  CGCAGTAATCATCATATGAGACCCCAAA 2188

RESULT 3
AX032762
LOCUS      3761 bp      DNA      linear      PAT 21-SEP-2000
DEFINITION      Sequence 2 from Patent WO046358.
ACCESSION      AX032762
VERSION      AX032762.1 GI:10279738
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
  Dean,C., West,J. and Johanson,U.
  Plant gene
  Patent: WO 0046358-A 2 10-AUG-2000;
  DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
  JOHANSON URBAN (SE)
  Location/Qualifiers
FEATURES

```

```

source      1..3761
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.:      9,02e-187      Length:      3761
Score:      2981.00      Matches:      608
Percent Similarity:      78.96%      Conservative:      0
Best Local Similarity:      78.96%      Mismatches:      1
Query Match:      96.04%      Indels:      162
DB:      2      Gaps:      2

US-09-890-475-1 (1-609) x AX032762 (1-3761)

Qy      1  MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu 20
Db      574  ATGTCCAAATATATCCACCGCGGTGGCGCAACCAACACGCGCAATCCATCTGCTG 633
Qy      21  GlnArgHisGlnSerGluGlnArgArgArgGluLeuProLysLileValGluThrGluSer 40
Db      634  CAGCGACATCATCTGACACGCGACGAGAGAAATACCGAAGATTGTGAAACAGAGTCT 693
Qy      41  ThrSerMetAspLileThrLileGlyGlnSerLysGlnProGlnPheLeuLysSerLileAsp 60
Db      694  ACAAGTATGGACATATACGATCGTCAATCTAAGCAGCTCAATTTTGAATCCATAGAC 753
Qy      61  GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80
Db      754  GAATTAGCTGCGTTTTCAGTTGAGTGGAAACATTCAAACGCAATTCGATGATCTTCAG 813
Qy      81  LysHisLileGluSerLileGluAsnAlaLileAspSerLysLeuGluSerAsnGlyValVal 100
Db      814  AAGCACATCGAGTCAATCGAAACGCAATGATTCAAACTCGAGAGTAAACGCGCTGTC 873
Qy      101  LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnVal 120
Db      874  CTCGCGCGCGGCAACAAATATTTCCATCAGCGCATGTTATCGCTCCGCGGAACTATGA 933
Qy      121  SerValGluThrThrValThrValSerGlnProSerGlnGluLileValProGluThrSer 140
Db      934  TCTGTAGAACCAACCGCTCACTGTGAGCCCAACCGTCTCAGAGATTTGACCGGAGACGTG 993
Qy      141  AsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
Db      994  AATAAACCGGAGGCGGAGCTATGCTGTGATGTTGATGTAGCAAGGCTCTGCGTAAATAC 1053
Qy      161  LileTyrAlaAsnLileSerAspGlnAlaLysLeuMetGluGluLileProSerAlaLeuLys 180
Db      1054  ATATACCGAATATCTCTGATCAAGCTAAGTTAATGGAAGAGATTCCTTCAGCTTTGAAA 1113
Qy      181  LeuAlaLysGluProAlaLysPheValLeuAspCysLileGlyLysPheTyrLeuGlnGly 200
Db      1114  TTGGCCAAAGGAGCGAGCAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTACAAAGG 1173
Qy      201  ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLile 220
Db      1174  CGTAGACATTTACTTAAGATCGCTTATGAGCTTCGCGAGACAAAGTTTCGTTCTTATA 1233
Qy      221  LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysLileGluSerTyr 240
Db      1234  CTGGAGTCTTTTCTTCTAATGCTGATCGTGTAAAGGGAAGGTGAAGATTGAGAGTTGG 1293
Qy      241  LileLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGly 260
Db      1294  ATTAAGATGAGGCGGAGCGGCTGCTGTGCTGGAGGAAAGGTTGATGACTGAGGA 1353
Qy      261  GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
Db      1354  GGATTAGCTCGGCTGAGAAATGGATGCAAGGGTTTGCTTTTACTAGTTGCTGTTT 1413
Qy      281  GlyValProSerAsnPheArgSerThrAspLeuLeuAspLileArgMetSerGlySer 300

```



Db 1414 GGTGTTCTCTCAAACTTTAGGAGTACAGATTGCTGATTTGATTAAGGATGAGTGGTTGC 1473  
Qy 301 AsnGluLeuAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer----- 318  
Db 1474 AATGAGATTGCCGGTGTCTTTGAAGCGGTACAGTTCTCTTGTGCCCTATGGTCTCAGGTACC 1533  
Qy 318 ----- 318  
Db 1534 ATATTCTGTTCTCACTCGGTGAATTTCATTGCAAAAGGTGGTTCCTTTGTTGACATCATC 1593  
Qy 318 ----- 318  
Db 1594 GACCAACATCAAGTTCACATCTTTGTTTTGATGAAGCTTGATGATATAACTAGGAGAGC 1653  
Qy 318 ----- 318  
Db 1654 ACATCAATATTAGAGTGCAATGACTGATTTAGCCAAATCCTAGCTAGAAAATTAATCTG 1713  
Qy 318 ----- 318  
Db 1714 GAAAGAACTTGGAACTCTCAACCATAGGTTTTGGTACGAAATTTGTTGTCAGAACCA 1773  
Qy 318 ----- 318  
Db 1774 AATGATAGGCTATTGCCTTGAATAGTGTCTTGTGGTTCCATATTGGAAGCTTAAA 1833  
Qy 318 ----- 318  
Db 1834 TCGTATGACTTAGCTTTGGATACCTAATAAGCTTAAGCAATGCCAACTCTTAAGAGTGG 1893  
Qy 319 -----GlyileValGluSerSerIleLysArgGlyMet 329  
Db 1894 TACTTACAAATATCTATTGTTGGTCATAGTATAGTTGAATCAAGTATCAGCGTGGATG 1953  
Qy 330 HisLeuAlaLeuGluMetValThrPheGlyMetGluAspLysPheSerAlaAla 349  
Db 1954 CATATTGAAGCTCTTGAGATGTTTATACCTTTGGCATGGAGGATAAGTTTTCAGCTGCT 2013  
Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369  
Db 2014 CTAGTTCTTAATCTTATTAAAGATGAGCAAGAGTCAATTTGAGAGGGCAAAACGGAAA 2073  
Qy 370 AlaGlnSerProLeuAlaPheLys----- 377  
Db 2074 GCCAGTACCCTGGCCATTTGT-ATGAACCTTCCCTTGCACATTATGTACCTTTATGA 2132  
Qy 378 -----GluAla 379  
Db 2133 ACTCTTTATCATCTCTGAGTCTGACCATTCATATATTTATTTCTCAACAGAAAGAGCG 2192  
Qy 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399  
Db 2193 GCTACAAAGCAGCTAGCTGTGTATCATCATGTTATGCACTGTATGGAGACTCACAGTTA 2252  
Qy 400 AspProAlaLysGluLeuProGlyTrpGlnLysGluGlnLysValSerLeuGluLys 419  
Db 2253 GATCTCGCAAGAACTACCAAGGATGGCAGATCAAGAGCAAAATTTGTTAGCTTCGAGAAA 2312  
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439  
Db 2313 GACACTCTTCAGCTCGACAAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTAATGAG 2372  
Qy 440 GluAlaLeuAlaLysArgMetTyrAsnGlnGlnLysArgProArgLeuSerPro 459  
Db 2373 GAAAGCCGCACTGCCAAGAGATGTATAACCAACAGATAAAACGTCACAGTTGTCAACC 2432  
Qy 460 MetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe 479  
Db 2433 ATGGAATATGCCACCACTTCTTCTCATCGTATTTCTCTTCTTCTTCTTCTTCTTCTTCTT 2492  
Qy 480 ProSerGlnArgAspAspGlnAspGlnLeuSerAlaLeuValSerSerTyrLeuGly 499

Db 2493 CCTAGTCAAAAGAGACGATGACCAAGATGAATATCAGCTCTTGTGAGTAGTACTCTCGC 2552  
Qy 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519  
Db 2553 CGGTCAACATCTTTTCTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCA 2612  
Qy 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539  
Db 2613 CATGTGGGTAGGAAGATGTATGATATGAACATCTGCCCCCAAAATTCATACTCT 2672  
Qy 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559  
Db 2673 CAGGTCACGACATAGACTTCATCGACAGTACTCTCGTCTTTGTTCCGACAGAGA 2732  
Qy 560 HisProLeuGlnTyrSerProIleHisGlyGlnGlnLeuProTyrGlyIleGln 579  
Db 2733 CATCCACTACAGTACTCTCTCTCAATTCATGACACACACAGTTTACCATATGTTATACA 2792  
Qy 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLysSerAsnGlnArgSer 599  
Db 2793 AGGTTTACAGACATTCACATCTGAGNAGATATTTGGGTTTATCCAATCAAGGTTCT 2852  
Qy 600 ProArgSerAsnSerSerLeuAspProLys 609  
Db 2853 CCTCGAGTAACTCATCTAGACCCCAA 2882

RESULT 4  
AP228499  
LOCUS AP228499 3761 bp DNA linear PLN 14-OCT-2000  
DEFINITION Arabidopsis thaliana FRIGIDA (FRI) Gene, complete cds.  
ACCESSION AP228499  
VERSION AP228499.1 GI:10801173  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 3761)  
AUTHORS Johanson,U., West,J., Lister,C., Michaels,S., Amasino,R. and Dean,C.  
TITLE Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time  
JOURNAL Science 290 (5490), 344-347 (2000)  
MEDLINE 20485641  
PUBMED 11030654  
REFERENCE 2 (bases 1 to 3761)  
AUTHORS Johanson,U. and Lister,C.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich, Norfolk NR4 7UH, U.K.  
FEATURES  
source  
1..3761  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/cultivar="ecotype Hs1"  
/db\_xref="taxon:3702"  
/map="between m151 and m122"  
/genes="FRI"  
join(<574..1528,1922..2094,2184..>2885)  
/genes="FRI"  
/product="FRIGIDA"  
join(574..1528,1922..2094,2184..2885)  
/notes="late flowering gene"  
/codon\_start=1  
/product="FRIGIDA"  
/protein\_id="AAG23414.1"  
/db\_xref="GI:10801174"  
/translation="MSNIPTVAQPTTTANPLLQHQSEQRRELPKIVETESTSMD  
ITIGSKQFOFLKSIDELAASFVAVETFRKQFDLDQKHIESIENALDSKLESGVVIA

ARNNHFPMLSPNNVSVTTVTSQSOEIVPSTNPKPEGRMCMLCSKGLRKY  
 IYANISDAKMEBIPSAKIAKEPAKFAVLDICIGFYLOGREAFPTKSPMSAROVSL  
 LILNLFPLMPRGKVKYIESWIKDEAFAVAWKELMTGEGGIAAEKMDARGLLUL  
 VACRFVPSNFDLDDILITMSGNEIAGALKRSQPIVPMVSGIVESIKRGWHEIAL  
 EMVYTFGMEKFSAAVLVITFLKMSKSFERAKRAQSPFAKFAATKQLALVSSVMQ  
 CMETHKLDPAKELPWQIKQIVLEKDTLQDKEMEEKARSLMLBEERALAKRMNQ  
 QIKRPLSPMPMPVTSYSPYIYDRSPFPSPDDQDEISALVSSYLGSTSPFPHRS  
 RRSPPYVPLPHGGIGRGSVAYAYHLAPNSYSPGHGRLHRYQSPSLVHGQRHPLQYSP  
 PTHGQQQLPYGIQRYVRHSPSEERYVGLGSLNQRSPRSNSLDPK"

## ORIGIN

## Alignment Scores:

Pred. No.: 9,02e-187 Length: 3761  
 Score: 2981.00 Matches: 608  
 Percent Similarity: 78.96% Conservativity: 0  
 Best Local Similarity: 78.96% Mismatches: 1  
 Query Match: 96.04% Indels: 162  
 DB: 8 Gaps: 2

US-09-890-475-1 (1-609) x AF228499 (1-3761)

QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20  
 Db 574 ATGTCCTCAATATCCACCGACGGTGGCGGCAACCAACACGCGCAATCCACTGCTG 633  
 QY 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40  
 Db 634 CAGCGACATCAATCTGAAACAGCAGCAGAGAGAAATACCGAAGATTGTGAAACAGAGTCT 693  
 QY 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60  
 Db 694 ACAAGTATGACATTAACGATCGGTCAATCTAAGCAGCCTCAATTTTGAATCCATAGAC 753  
 QY 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspSerLeuGln 80  
 Db 754 GAATTAGCTCGCGTTTTCAGTTCAGTGGAAACATTCAACGCAATTCGATATCTTCAG 813  
 QY 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100  
 Db 814 AAGCACATCGATCAATCGAAACGCAATGTATCCAACTCGAGATTAACGCGTGTGC 873  
 QY 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120  
 Db 874 CTCGCGCGCGGAACAATAATTTCCATCAGCCGATGTTATCGCTCCGCGGAACAATGTA 933  
 QY 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140  
 Db 934 TCTGTAGAACACCCCTCACTGTGAGCCCAACCGTCTCAGAGATTGTACCGAGAGCTCG 993  
 QY 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160  
 Db 994 AATAAACCGGAGGGGAGCGTATGTGTGAGTTGATGTAGCAAGGCTCGCGTAATAATAC 1053  
 QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180  
 Db 1054 ATATACGCGAATATCTCTGATCAAGTAGATTAAATGGAAGAGATTCCTTCAGCTTTGAAA 1113  
 QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200  
 Db 1114 TTGCGCCAGAGAGCAGCAGAGTTTGTATTGATTGTTGGCAAGTTTACTTACAAAGG 1173  
 QY 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220  
 Db 1174 CGTAGAGCATTTACTAAAGAGTCGCTATGAGCTCTGCAGAGCAAGTTTCGCTTCTTATA 1233  
 QY 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrp 240  
 Db 1234 CTGGAGTCTTTTCTTAAATGCTGATCGTGGTAAAGGAAGGTGAAGATTGAGAGTTGG 1293  
 QY 241 IleLysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGly 260  
 Db 1294 ATTAAGAGATGAGCGGAGCGGTGCTGTGCTTGGAGGAAAAGGTTGATGACTGAAGGA 1353

QY 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280  
 Db 1354 GGATTAGCTGCGGCTGAGAAAAATGGATCAAGGGGTTGCTTTTATAGTTGCTTGT 1413  
 QY 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300  
 Db 1414 GGTGTTCCCTTCAACTTTAGGAGTACAGATTTCGTGGATTGTAAGGATGAGTGCTCG 1473  
 QY 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318  
 Db 1474 AATGAGATTGCGGCTGCTTTGAAGCGGTCACAGTTTCTTGTCCCTATGCTCTCAGGTACC 1533  
 QY 318 ----- 318  
 Db 1534 ATATTCTGTTCTCACTCGGTGAATTTTCATTGCAAGGTTGGTTCCTTTGTTGTCATCATC 1593  
 QY 318 ----- 318  
 Db 1594 GACCAACATCAAGTTCCTCTTTGTTTTTCGATAGCTTTGATGGTATATAAACTAGGAGAGC 1653  
 QY 318 ----- 318  
 Db 1654 ACATCAAAATATTTAGAGTGCATGACTGATTGAGCCAAATCCTAGCTAGAAATTAATCTG 1713  
 QY 318 ----- 318  
 Db 1714 GAAAGAACTTGGAACTCTCAACCATAGTTTTGGTACGAAATGTTGCTTGTTCAGAACCA 1773  
 QY 318 ----- 318  
 Db 1774 AATGATAGGCTATTGCGCTTGAANAATAGTGTTCCTTTGTTGGTTTCCAAATATTGGAAGTTAAA 1833  
 QY 318 ----- 318  
 Db 1834 TCGTATGACTTAGCTGTGGTACTAATAAGCTTAAGCAATGCCAACTCTAAGAAAGTGG 1893  
 QY 319 ----- GlyIleValGluSerSerIleLysArgGlyMet 329  
 Db 1894 TACTTACACAATATTCATTGTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1953  
 QY 330 HistGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349  
 Db 1954 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGGATAAGTTTTCAGCTGCT 2013  
 QY 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369  
 Db 2014 CTAGTTCTAACTTCAATCTTTAAAGATGAGCAAGAGTCAATTTGAGAGGGCAAAACGAAA 2073  
 QY 370 AlaGlnSerProLeuAlaPheLys ----- 377  
 Db 2074 GCCCAGTCACCGCTGGCATTTTGT-ATGAACCCCTTCCCTTGCACATTATGTACCTTTATGA 2132  
 QY 378 ----- GluAla 379  
 Db 2133 ACTCTTTATCATCATCTGAGTCTGACCATTTGATATATTTATTTCTCAACAGAAAGAGCG 2192  
 QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399  
 Db 2193 GCTCAAGAGCAGCTAGCTGTGTTATCATCATGATTATGAGTGTATGAGACTCACAGTTA 2252  
 QY 400 AspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLys 419  
 Db 2253 GATCCTCGAAAGAACTACCAAGGATGGCAGATCAAGAGCAAAATTTGTTAGCTTGGAGAAA 2312  
 QY 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439  
 Db 2313 GACACTCTTCAGCTCGCAAAAGAGATGGAAGAGAAACAGATCTCTCAGTTTATGGAG 2372  
 QY 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459  
 Db 2373 GAAGCGGAGCTTGCACAGAGATGTATAACCAACAGATAAAACGTCACAGGTTCTCACCC 2432  
 QY 460 MetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSerPhe 479



```
Db 1114 ATTAAGATGAGCGGAGAGCGCTGCTGTTGCTTGGAGAAAGGTTGATGACTGAAGGA 1173
Qy 261 GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
Db 1174 GGATTAGCTCGGCTGAGAAAATGGATGCAAGGGTTTGCTTTACTAGTTGCTTGTTTT 1233
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuArgMetSerGlySer 300
Db 1234 GGTGTTCTTCAAACTTTAGGAGTACAGATTTGCTGGATTTGATGAAGATGATGGTTCG 1293
Qy 301 AsnGluIleAlaGlyValAlaLysArgSerGlnPheLeuValProMetValSer 318
Db 1294 AATGAGATTGCCGGTGCTTTGAAGCGGTACAGTTTCTTGCCCTATGGTCTCAGGTACC 1353
Qy 318 ----- 318
Db 1354 ATATTCTGTTCTCACTCGGTGAATTTCAATGCAAAAGGTGGTCTCTTTGTTGACATCATC 1413
Qy 318 ----- 318
Db 1414 GACCAACATCAAGTCCATCTTTGTTTTCGATAAGCTTGATGGTATAAACTAGGAGAGC 1473
Qy 318 ----- 318
Db 1474 ACATCAAAATATTAGAGTGCAATGACTGATTGAGCCAAATCCTAGCTAGAAATTAATCTG 1533
Qy 318 ----- 318
Db 1534 GAAAGAACTTGGAACTCAACCATAGGTTTGGTACGAAATTTGTTGTCAGAACCA 1593
Qy 318 ----- 318
Db 1594 AATGATAGGCTATTGCCITTTGAAATAGTGTCTTCTGTGTTTCCAATATTGGAAGTAAAA 1653
Qy 318 ----- 318
Db 1654 TCATATGACTTAGCTTGGATACTAATTAAGCTTAAGCAATGCCCACTCAAGAAAGTGG 1713
Qy 319 -----GlyIleValGluSerSerIleLysArgGlyMet 329
Db 1714 TACTTACACAATATTCTATTGGTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1773
Qy 330 HisIleGluAlaLeuGluMetValThrPheGlyMetGluAspLysPheSerAlaAla 349
Db 1774 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGATGAAGTTTTCAGCTGCT 1833
Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
Db 1834 CTAGTTCTTAACCTTCACTTTAAAGATGAGCAAGGAGTCATTTGAGAGGCAAAACGGAAA 1893
Qy 370 AlaGlnSerProLeuAlaPheLys----- 377
Db 1894 GCCAGTCACCGCTGGCATTTGT-ATGAACCTTCCCTTGCACATTATGATCCTTTATGA 1952
Qy 378 -----GluAla 379
Db 1953 ACTCTTTATCATCTGAGTCTGACCATTGATATATTATTCTCAACAGAAAGAGCG 2012
Qy 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399
Db 2013 GCTACAAAGCAGCTAGCTGTGTTTATCATCAGTTATGCAAGTATGATGGAGACTCACAAGTTA 2072
Qy 400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLys 419
Db 2073 GATCCTCGCAAGAACTACCAAGATGCGAGATCAAGAGCAAAATGTTAGCTTGGAGAAA 2132
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439
Db 2133 GACACTCTTCAGCTCAGCAAAAGAGATGGAAGAGAAACCAAGATCTCTCAGTTTAAATGGAG 2192
Qy 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459
```

```
Db 2193 GAAGCGCAGCTTCCCAAGAGATGTATATAACCAACAGATAAAACGTCACAGGTTGTCAACC 2252
Qy 460 MetGluMetProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe 479
Db 2253 ATGGAATATGCCACCACTAACTTCTTCATCGTATTTCTTCATCTACCGTGATGAAGCTTT 2312
Qy 480 ProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGly 499
Db 2313 CCTAGTCAAGAGACCATGACCAAGATGAATATACAGCTCTTGAGTAGTAGTTACCTCGGC 2372
Qy 500 ProSerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProLeuPro 519
Db 2373 CCCTCAACATCTTTCTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCA 2432
Qy 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
Db 2433 CATGGTGGTTAGGAAGAGTGTATATGCATATGAACATCTGGCCCAAAATTCATACTCT 2492
Qy 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
Db 2493 CCAGGTCAACGACATAGACTTCATCGACAGTACTCTCCGTCTTTGGTTCCACGACAGAGA 2552
Qy 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnGlnLeuProTyrGlyIleGln 579
Db 2553 CATCCACTACGACTCTCTCTCCAAATTCATGCAACAACAGATTACCATATGGTATACAA 2612
Qy 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
Db 2613 AGGGTTTACAGACATTCACCATCTGAAGAAGATATTGGGTTTATCCCAATCAAGGTCT 2672
Qy 600 ProArgSerAsnSerSerLeuAspProLys 609
Db 2673 CCTCGCAGTAACCTCATCTAGTACCCCAAA 2702

RESULT 6
AV198402
LOCUS
DEFINITION Arabidopsis thaliana ecotype KZ-9 FRIGIDA (FRI) gene, complete cds.
ACCESSION AV198402
VERSION AV198402.1 GI:31558912
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
REFERENCE
AUTHORS Gazzani,S., Gendall,A.R., Lister,C. and Dean,C.
TITLE Analysis of the molecular basis of flowering time variation in Arabidopsis accessions
JOURNAL Plant Physiol. 132 (2), 1107-1114 (2003)
AUTHORS
TITLE 2 (bases 1 to 3261)
JOURNAL Plant Physiol. 132 (2), 1107-1114 (2003)
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes Center, Colney Lane, Norwich NR4 7UH, UK
FEATURES
source
1..3261
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="4"
/map="between mi51 and mi122"
/ecotype="KZ-9"
<505..>2816
/gene="FRI"
join(<505..1459,1853..2025,2115..>2816)
/product="FRIGIDA"
join(505..1459,1853..2025,2115..2816)
/gene="FRI"
/note="flowering gene"
/codon_start=1
```

```
/product="PRIGIDA"  
/protein_id="AAP49808.1"  
/db_xref="GI:31558913"  
/translation="MSNPFPIVAAQPTTANPLLQHQSQRRRELPKIVETSTSD  
ITIGSQPQLKSIDLAFAVETFKQFDDLQKHIESIENAIIDSKESNGVIA  
ANNFNHQPMLSPRNVSVEITVVSQOEIVPETSNGPEGRMELMCSGLRKY  
IYANISQAKMEIPKALKAPEKFLVDCIGFYLOQRAPFTKESPMSSARQVSL  
LILESFLIMPDRGKVKIESWIKDEAETAAVWRKELMTTEGGIILAAEKMDARGLLIL  
VACFGVSNPSTDLILLRMSGNEIAGLRKSOFLVPMVSGIVGESIKRGVHIBAL  
EMVTFQWEDKFSALVLTSPFKMSKESFERAKKAQSPAPKFAAATKQIAVLSSVYQ  
CMETHKLDPAKELFGWQIKQIVSLEKDTQLDKEMEKEKARSLSLMEALAKRMVQ  
QIKRPLSPMPGVPTSSSYPIYRDSFPQDDQDEISALVSSVLPFGSTFPFRS  
RRSPYVPLPGLGHSVAYEHLAPNSPSYGHGHLRQYSPSLVHGHRPLQVSP  
PIHQQQULPYGIQRYVRHSPSEERYLGLSNQSPRSNSLDPK"
```

## ORIGIN

```
Alignment Scores:  
Pred. No.: 1,866-186 Length: 3261  
Score: 2975.00 Matches: 607  
Percent Similarity: 78.83% Conservative: 0  
Best Local Similarity: 78.83% Mismatches: 2  
Query Match: 95.84% Indels: 162  
DB: 8 Gaps: 2
```

US-09-890-475-1 (1-609) x AY198402 (1-3261)

```
Qy 1 MetSerAsnTyrProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20  
Db 505 ATGTCCAATTATCCACCACGGTGGCGCGCAACCCCAACAGCGCGGAATCCACTGCTG 564  
Qy 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40  
Db 565 CAGGACATCAATCTGAACACGACGACGACGAGAAATTCGGAAGATGTGCGAACAAGATCT 624  
Qy 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60  
Db 625 ACAAGTATGGACATTACGATCGGTCAATCTAAGCAGCGCTCAAAATTTGAAATCCATAGAC 684  
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80  
Db 685 GAATTAGCTGGCTTTTCAGTTGCGAGTGGAAACATTTCAACGCCAAATTCGATGATCTTCAG 744  
Qy 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100  
Db 745 AAGCACATCGAGTCAATCGAAACGCAATTGATTCGAACTCGAGAGTAAACGGCGTGTGC 804  
Qy 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120  
Db 805 CTCGCCGCGCGGAAACAATAATTTCCATCAGCCGATGTTATCGCTCCGCGGGAACAATGTA 864  
Qy 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140  
Db 865 TCTGTAGAAACCAACCGTCACTGTGAGCCACCGTCTCAGGAGATTTGACCGGAGACGCTCG 924  
Qy 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160  
Db 925 AATAAACCGGGGGGGGACGATGTGTGTGAGTGTGATGTGTAGCAAGGCTGCGTAAATAC 984  
Qy 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180  
Db 985 ATATACGCCAATATCTCTGATCAAGCTAAGTTAATGGNAGAGATTCCTTCAGCTTTGAAA 1044  
Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200  
Db 1045 TTGGCCAGGAGGACGAGGAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1104  
Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220  
Db 1105 CGTAGAGCATTTACTAAAGATGTCGCTATGAGCTCTCGGACAAAGTTTCGCTTCCTATA 1164  
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTyr 240  
Db 1165 CTGGAGTCTTTTCTTAATGCTGATCGTGGTAAAGGAGGATGAAGATTGAGAGTTGG 1224
```

```
Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaIleTyrArgLysArgLeuMetThrGluGly 260  
Db 1225 ATTAAGATGAGCGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1284  
Qy 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPhe 280  
Db 1285 GGATTAGCTGCGGCTGAGAAAATGATGCAAGGGGTTGCTTTTACTAGTTCTGCTTTT 1344  
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300  
Db 1345 GGTGTTCTTCAAACTTTAGGAGTACAGATTTGCTGGATTTGATAGGATGAGTGGTTCG 1404  
Qy 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318  
Db 1405 AATGAGATTGCGGCTGCTTTGAAGCGGTCAAGTTTCTTGTCCCTATGCTCAGGTACC 1464  
Qy 318 ----- 318  
Db 1465 ATATTCGTCTCACTCGGTGAATTTCAATGCAAAAGTGGTTCCTTTGTTGACATCATC 1524  
Qy 318 ----- 318  
Db 1525 GACCAACATCAAGTTCACATCTTTGTTTTCGATAAGCTTGATGTTATAAACTAGGAGAGC 1584  
Qy 318 ----- 318  
Db 1585 ACATCAAAATATTAGAGTGCAATGACTGATTGAGCAAAATCTAGCTAGAAATTAATCTG 1644  
Qy 318 ----- 318  
Db 1645 GAAAGAACTTGGAACTCTCAACCATAGGTTTTGGTACGAAATTTGTTGTTGTCAGAACCA 1704  
Qy 318 ----- 318  
Db 1705 AATGATAGGCTATTGCTTTGAAATAGTGTCTTCTGTGGTTTCCAATATTTGGAAGTTAAA 1764  
Qy 318 ----- 318  
Db 1765 TCATATGACTTAGCTTTGGATACTAATTAAGCTTAAGCAATGCCAATCTTAAGAAGTGG 1824  
Qy 319 -----GlyIleValGluSerSerIleLysArgGlyMet 329  
Db 1825 TACTTACACAAATATTCTATTGCTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1884  
Qy 330 HisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349  
Db 1885 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGATAAGTTTTCAGCTGCT 1944  
Qy 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369  
Db 1945 CTAGTTCTTAATCTTCAATCTTAAAGATGAGCAAGGAGTCAATTTGAGAGGGCAAAACGGAAA 2004  
Qy 370 AlaGlnSerProLeuAlaPheLys ----- 377  
Db 2005 GCCCAGTCCACCTGCGCATTTGT-ATGAACCCCTCCCTTGCACATTATGTACCTTTATGA 2063  
Qy 378 -----GluAla 379  
Db 2064 ACTCTTTATCATCATCTGAGTCTGACCATTTGATATATTATTCTCAACAGAAAGACG 2123  
Qy 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399  
Db 2124 GGTACAAACAGCTAGCTGTGTATCATCAGTTATGCGATGATGAGACTCACAAGTTA 2183  
Qy 400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGluIleValSerLeuGluLys 419  
Db 2184 GATCTCTGCAAGAACTACCAAGTGGCGAGTCAAGAGCAAAATTTGTAGCTTTGAGAAA 2243  
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439  
Db 2244 GACACTCTTCAGTCCGCAAAAGAGATGGAAGAAAGCAAGATCTCTCAGATTTAATGGAG 2303
```

```
QY 440 GluAlaLeuAlaLeuArgMetTyrAsnGlnGlnIleValArgProArgLeuSerPro 459
DB 2304 GAAGCGGCACITGCCAGAGATGATTAACCAACAGATAAAAGTCCAGGTTCTACCC 2363
QY 460 MetGluMetProProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPhe 479
DB 2364 ATGGAATGCCACCAAGTAATCTTCTATCGTATCTCTCTATCTACCGTGATAGAAGCTTT 2423
QY 480 ProSerGlnArgAspAspGlnAspGlnIleSerAlaLeuValSerSerTyrLeuGly 499
DB 2424 CCTAGTCAAGAGAGAGATGACCAAGATGAATATCATGCTCTTGAGTAGTACTCGGC 2483
QY 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519
DB 2484 CGCTCAACATCTTTCTCATCGCTCAAGAAGATCCCGGAATATATGGTTCCACTTCCA 2543
QY 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
DB 2544 CATGTGGGTAGGAAGAGTGTATGATCATATGAACATCTGGCCCCAAATTCATATCTCT 2603
QY 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
DB 2604 CCAGTCCAGGACATAGACTTCATCGACAGTACTCTCGTCTTTGGTTTCACGGACAGAGA 2663
QY 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnGlnLeuProTyrGlyIleGln 579
DB 2664 CATCCACTACAGTACTCTCTCCATTCATGGACAAACAGTTACCATATGGTATACAA 2723
QY 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
DB 2724 AGGTTTACAGCATTCACCTCTGAAGAAGATATTGGTTTATCCATCAAGGTCT 2783
QY 600 ProArgSerAsnSerSerLeuAspProLys 609
DB 2784 CCTCGCAGTAATCATCATATTAGACCCCAAA 2813

RESULT 7
AY198401
LOCUS AY198401 3739 bp DNA linear PLN 12-JUN-2003
DEFINITION Arabidopsis thaliana ecotype Shakhara FRIGIDA (FRI) gene, complete cds.
ACCESSION AY198401
VERSION AY198401.1 GI:31558910
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
1 (bases 1 to 3739)
Gazzani,S., Gendall,A.R., Lister,C. and Dean,C.
Analysis of the molecular basis of flowering time variation in
Arabidopsis accessions
Plant Physiol. 132(2), 1107-1114 (2003)
2 (bases 1 to 3739)
Lister,C. and Dean,C.
Direct Submission
Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes
Center, Colney Lane, Norwich NR4 7UH, UK
FEATURES
Location/Qualifiers
1..3739
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="4"
/maps="between mi51 and mi122"
/ecotype="Shakhara"
<574..>2885
/gene="FRI"
join(<574..1528,1922..2094,2184..>2885)
/gene="FRI"
/product="FRIGIDA"
join(574..1528,1922..2094,2184..2885)
CDS
```

```
/gene="FRI"
/note="flowering gene"
/codon_start=1
/product="FRIGIDA"
/protein_id="AAP49807.1"
/db_xref="GI:31558911"
/translation="MSNYPTVAQPTTANPLLQHQSEORRELPKIVFTSTMSD
ITIGSKQPOLKSIDELAASFVAETFKROFDDLOKHIESIENADSKLESNGVLLA
ANNNNFOPMLSPNNVSVETVTVSQSDEIVPETSNGKRGRAFTKESPMASARQVSL
IVANISDAQKMEEPISALKLAKPAKVLDCIGKFLQGRRAFTKESPMASARQVSL
LILESFLAMPDRGKGVKIESWIKDEAETAAVAMFKLMTGGGLAAAEKMDARGQLLL
ENYFVGMEDKFSALVLTSFKMSKESFERAKRQAQSPFLAKKEATKQLAVLSSVMQ
CMETKFLDPKAPKQWIKQIYSLEKDTLQDKEMEEKARSLSMEEAALAKRMYNQ
QIKRKLSPMEMPPVTSSYSIYRDRSPQSDDDDEISALVSYLGPTSPFQRS
RRSPYMWPLPHGGLGRSVYAEHLAPNSYSPGHGRLHRYQSPSLVHGQRHPLYSP
PIHGQQLPYGIQVRVHRSPSEERYVLGLSNQSRSPSNSSLDPK"
ORIGIN
Alignment Scores:
Pred. No.: 2,226-186 Length: 3739
Score: 2975.00 Matches: 607
Percent Similarity: 78.83% Conservative: 0
Best Local Similarity: 78.83% Mismatches: 2
Query Match: 95.84% Indels: 162
DB: 8 Gaps: 2
US-09-890-475-1 (1-609) x AY198401 (1-3739)
QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu 20
DB 574 ATGTCCATTAATCCACGAGCGTGGCGCGCAACCAACGACGCGAATCCACTGCTG 633
QY 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40
DB 634 CAGCGACATCAATCTGAACGCGACGAGAGAAATACCGAAGATTGTGAAACAGAGTCT 693
QY 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60
DB 694 ACAAGTATGGACATTAACGATCGTCAATCTAAGCAGCCTCAATTTTGAATCCATAGAC 753
QY 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspLeuGln 80
DB 754 GAATTAGTTCGCTTTTTCAGTTGAGTGGAAACATTCACAAACGCGCAATTCATGATCTTCAG 813
QY 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100
DB 814 AAGCATCATGATCAATCGAAGAACGCAATGATTCGAACTCGAGAGTAACGCGCTTGTC 873
QY 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120
DB 874 CTCGCGCGCGGAACCAATAATTTCCATCAGCGATGTTATCGCTCCGCGGAACATGTA 933
QY 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140
DB 934 TCTGTAGAAACCAACCGTCACTGTGAGCCACCGTCTCAGAGATTTGATCCGAGAGAGCTG 993
QY 141 AsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
DB 994 AATAAACCGGAGGGGAGCGATGTTGTGAGTTGATGTTAGCAAGGTTCTGCGTAATATAC 1053
QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
DB 1054 ATATACGCGAATATCTCTATCAAGCTAAGTTAATGGAAGAGATTCCTTCAGCTTTGAAA 1113
QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
DB 1114 TTGCGCCAGGAGCCAGCGAAGTTTGTATTGGATTGTTATGGCAAGTTTCTTCAGCTTTG 1173
QY 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLys 220
DB 1174 CGTAGAGCATTTACTAAAGAGTCGCTATGAGTCTCTGCGAGACAAAGTTTTCGCTTCTATA 1233
```





```
/gene="PRI"
/product="FRIGIDA"
Join(527..1481,1875..2047,2137..2838)
/gene="PRI"
/note="flowering gene"
/codon_start=1
/product="FRIGIDA"
/protein_id="AAP49810.1"
/db_xref="GI:3155818"
/translation="MSNYPPTVAQPTTANPLQRHOSORRELPKIVETESTMD
ITIGQSQFQLSKDELAFSAVETFKQDFDLQKHIESIENAIKSKLESGVILA
ARNNHFOPLSPNNNTVETTVTSQFSEIVPETSNNKEGGRMCMLMSKGLRY
IVANTSEOKAMEEIPSAIKLAKEPAKFLVDCIGFYQGRRAFTKSPMSARQVSL
LILBSFLMPDRGKGVKIESWIKDEASTAAVAKRLMTTEGGLAAAEKMDARGLLIL
VACFGVPSFSTDLDLIRMSGNEIAGALKRSQFLVPMVSGIVESIKEGMHEAL
ENVYTFGMEDEKFSALVILTSFLKSKSEPERAKRKAQSPFLAPKAAKQLAVLSVMQ
QMKTHLPAKELPMQIKQIVLSLEKDTLQDKMEBEKARSLMSLEAAKAKMNYQ
QIKRPLRLEMPPTSSYPIYDRSPFSDDDODEISALVSSYLGPSSTSPHRS
RSPRYMPLPHGGLGRSVAYEHLAPNSYSPFGHRLHROYSPLVHGQRHPLYSP
PIHQOQLPYGIQRYRHSPEERYLGLSNRSPRSNSSLDPK"
ORIGIN
Alignment Scores:
Pred. No.: 4,698-186 Length: 3301
Score: 2969.00 Matches: 606
Percent Similarity: 78.83% Conservative: 1
Best Local Similarity: 78.70% Mismatches: 2
Query Match: 95.65% Indels: 162
DB: 8 Gaps: 2
US-09-890-475-1 (1-609) x AY198405 (1-3301)
QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20
Db 527 ATGTCCAAATATCCACCGACGGTGGCGGCCAACCCACACACGACGGCAATCCACTGCTG 586
QY 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLeuValValGluThrGluSer 40
Db 587 CAGCGACATCAATCTGAACGACGACGACGACGACGACGACGACGACGACGACGACGAC 646
QY 41 ThrSerMetAspIleThrIleGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60
Db 647 ACAAGTATGCACATACGATCGGTCAATCTAAGCAGCGCTCAATTTTGAATCCATAGAC 706
QY 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspLeuGln 80
Db 707 GAATTAGTCGCGTTTTCAGTTCAGTGGAAACATTCAAATGCCAATTCGATGATCTTCAG 766
QY 81 LysHisIleGluSerIleGluAlaIleAspSerLysLeuGluSerAsnGlyValVal 100
Db 767 AGGCACATCGAGTCAATCGAAGCAATGATTCCAACTCGAGAGTAACGCGCTTCTC 826
QY 101 LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal 120
Db 827 CTCGCCGCGCGAACAATAATTTCCATCAGCGCGATGTTATCGCTCCCGGCAACAATGTA 886
QY 121 SerValGluThrThrValThrValSerGlnProSerGlnIleValProGluThrSer 140
Db 887 TCTGTAGAAACCAACCGTCTACTGTAGCCCAACCGTCTCAGAGATTGTACCGGAGACGTG 946
QY 141 AsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
Db 947 AATAAACCGGAGGGGAGGTATGTGTGAGTTGATGTGTAGCAAGGTCGTGCTAAATAC 1006
QY 161 IleTyrAlaAsnIleSerGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
Db 1007 ATATACCGCAATATCTCTGAACAGCTAAGTTAATGAAGAGATTCTTTCAGCTTTGAA 1066
QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
Db 1067 TTGGCCCAAGGAGCGGAGGAGTTGTGATTCGATTTGATTGGCAAGTTTACTTACAAGGG 1126
QY 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220
1127 CGTAGAGCAATTTACTAAAGAGTCGCTATGAGCTCTCGAGACAAGATTTCGCTTTATA 1186
221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyValLysIleGluSerTyr 240
1187 CTGGAGTCTTTTCTTCTAATGCTGATCGTGTGTAAGGGAAGGTAAGATTGAGAGTTGG 1246
241 IleLysAspGluAlaGluThrAlaAlaValAlaTTPAtqLysArgLeuMetThrGluGly 260
1247 ATTAAGATGAGCGGAGACCGCTGCTGCTTGTGAGGAGAAAGGTTGATGATGAGGA 1306
261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
1307 GGATTAGTCGCGCTCAGAAAAATGGATGCAAGGGGTTGCTTTACTAGTTGCTTTT 1366
281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300
1367 GGTGTTCTCTTCAACCTTATGAGTACAGATTGCTGATTTGATGAGGAGTATGATGCTG 1426
301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318
1427 AATGAGATTGCGCGTCTTGAAGCGTCAAGTTTCTTGTCCCTATGCTCTCAGGTACC 1486
318 ----- 318
1487 ATATCTGTTCTCCTCGGTGAATTCATTGCAAGGTGGTCTCTTTGTTGACATCATC 1546
318 ----- 318
1547 GACCAACATCAAGTTCATCTTTTTCGATTAAGCTTGTGATATAAACTAGGAGAGC 1606
318 ----- 318
1607 ACATCAAAATATTTAGAGTGCAATGACTGATGAGCCAAATCTCTAGCTAGAAATTAATCTG 1666
318 ----- 318
1667 GAAAGAACTTGGAACTCTCAACCATAGTTTGGTAGCAAAATGTTGCTTGTGAGAACCA 1726
318 ----- 318
1727 AATGATAGGCTATTGCTTGAATAGTGTCTTGTGTTTCCATATATGGAAGTTAAAA 1786
318 ----- 318
1787 TCATGACTTAGCTGTGATTAATTAAGCTTAAGCAATCCCAACTTAGAAGTGG 1846
319 ----- GlyIleValGluSerSerIleLysArgGlyMet 329
1847 TACTTACACAATATTCTATTGGTCATAGGTATAGTTGAATCAAGTATCAAGCGTGAATG 1906
330 HisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349
1907 CATATTGAAGCTCTTGAGATGGTTTATACCTTTGGCATGGAGGATAGTTTTCAGCTGCT 1966
350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
1967 CTAGTTCTTAACCTTCACTTTTAAAGATGAGCAAGAGTCACTTTGAGAGGGCAAAACGAAA 2026
370 AlaGlnSerProLeuAlaPheLys ----- 377
2027 GCCAGTCACCGCTGGCATTGT-ATGAACCTTCCCTTGCACATTATGATGACCTTTATGA 2085
378 ----- GluAla 379
2086 ACTCTTTTATCATCATCTGAGTCTGACCATTTGATATATTTTCTCAACAGAAAGAGCG 2145
380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399
2146 GCTACAAAGCAGCTAGCTGTGTTATCATCAGTTATGAGTGTATGAGACTTCAAGATTA 2205
400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGluIleValSerLeuGluLys 419
```



```
Db 2206 GATCTCTGCGAAGAACTACCGAGATGGCAGATCAAGAGCAAATGTTAGCTGGAGAAA 2265
Qy 420 AspThrLeuGlnLeuAspLysGluMetGluLysAlaArgSerLeuSerLeuMetGlu 439
Db 2266 GACACTCTTCAGCTCGCAAGAGATGGAAGAGAAAGCAAGATCTCTCAGTTTAATGGAG 2325
Qy 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459
Db 2326 GAAGCCGACATGGCCAGAGAAATGTATACCAACAGATAAAACGTCCAAGGTGTTCACCC 2385
Qy 460 MetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSerPhe 479
Db 2386 ATGGAAATGCCACCACTTAATCTTCATCGTATCTCTCTCTCTCTCTCTCTCTCTCTCT 2445
Qy 480 ProSerGlnArgAspAspGlnAspGlnLysSerAlaLeuValSerSerTyrLeuGly 499
Db 2446 CCTAGTCAAGAGACGATGACCAAGATGAAATATCAGCTCTGTGGAGTAGTTACCTCGGC 2505
Qy 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519
Db 2506 CCGTCAACATCTTTCTCATCGCTCAGAGAGATCCCGGATATATGTTCCACTTCCA 2565
Qy 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
Db 2566 CATGCTGGGTAGGAAGAAGTGTATGTCATATGAACATCTGGGCCCAAAATTCATACTCT 2625
Qy 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
Db 2626 CCAGTTCAGGACATAGACTTCATCGACAGTACTCTCGTCTTTGGTTCCAGACAGAGA 2685
Qy 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnLeuProTyrGlyLeuGln 579
Db 2686 CATCCACTACAGTACTCTCCCTCCAATTCATGACAAACAGTATACCATATGGTATACAA 2745
Qy 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
Db 2746 AGGGTTACAGACATTCACCATCTGGAAGAAGATTTTGGTTATTCATCAATCAAAGTCT 2805
Qy 600 ProArgSerAsnSerSerLeuAspProlys 609
Db 2806 CCTCGCAGTAATCATCATTAGACCCCAAA 2835

RESULT 9
AY198404
LOCUS
DEFINITION
Arabidopsis thaliana ecotype Cvi FRIGIDA-like protein gene,
complete sequence.
ACCESSION
AY198404
VERSION
AY198404.1 GI:31558916
KEYWORDS
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 3758)
Gazzani, S., Gendall, A.R., Lister, C. and Dean, C.
Analysis of the molecular basis of flowering time variation in
Arabidopsis accessions
Plant Physiol. 132 (2), 1107-1114 (2003)
2 (bases 1 to 3758)
Gazzani, S. and Dean, C.
Direct Submission
Submitted (17-DEC-2002) Cell and Developmental Biology, John Innes
Center, Colney Lane, Norwich NR4 7UH, UK
Location/Qualifiers
1. 3758
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="4"
/map="between mi51 and mi122"
/ecotype="Cvi"
FEATURES
source
```

```
misc_feature 577..2888
/Note="similar to FRIGIDA (FRI)"
ORIGIN
Alignment Scores: 9.92e-185 Length: 3758
Pred. No.: 2950.00 Matches: 604
Score: 78.5% Conservative: 1
Best Local Similarity: 78.44% Mismatches: 4
Query Match: 95.04% Indels: 162
DB: 8 Gaps: 2
US-09-890-475-1 (1-609) x AY198404 (1-3758)
Qy 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20
Db 577 ATGTCCAAATTTCCACCGACGTTGGCGGCAAAACCAACAGCCGCGAATCCACTGCTG 636
Qy 21 GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer 40
Db 637 CAGCGACATCAATCTGAACAGCGACGAGAGAAATTACCGAAGATTGTCGAAACAGAGTCT 696
Qy 41 ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp 60
Db 697 ACAAGTATGGACATTACGATCGTCAATCTAAGCAGCCCTCAATTTTGAATCCATAGAC 756
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80
Db 757 GAATAGCTCGCTTTTCAGTTCAGTGGAAACATTCAATGCCAATTCGATGATCTTCAG 816
Qy 81 LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal 100
Db 817 AAGCACATCGAGTCAATCGAAACCGCAATTGATTCAAACTCGAGAGTAACCGCGTTGTC 876
Qy 101 LeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProArgAsnAsnVal 120
Db 877 CTGCGCGCGGGAACAAATAATTTTCATCAGCCGATGTTATCGCTCCGCGGAAACAATGTA 936
Qy 121 SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer 140
Db 937 TCTGTAGAAACCAACCGTCTGTGAGCAACCGTCTCAGAGATTGTACCGAGAGCGTGC 996
Qy 141 AsnLysProGluGlyValArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160
Db 997 AATAAACCGGAGCGGAGCGATGTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1056
Qy 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180
Db 1057 ATATACGCGAATATCTCTGCAACAGCTAAGTTAATGGAAGAGATTCCTTCAGCTTTGAAA 1116
Qy 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200
Db 1117 TTGGCCCAAGGAGCCAGCAAGTTTGTATTGGATTGTATTGGCAAGTTTACTTACAAGGG 1176
Qy 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle 220
Db 1177 CGTAGAGCAATTTACTAAAGATCGCTCTGCGAGACAGTTCGCTTCTTATA 1236
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTyr 240
Db 1237 CTGGAGTCTTTTCTTCTAATGCTCTGATCGTGTGTTAAGGGAAGGTGAAGATTGAGAGTGG 1296
Qy 241 IleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGly 260
Db 1297 ATTTAAGATGAGCGGAGACGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1356
Qy 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280
Db 1357 GGATTAGTCTCGGCTCAGAAAAATGGATGCAAGGGGTTTCTTTTACTAGTTGCTGTTTT 1416
Qy 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300
Db 1417 GGTGTTCCCTCAAACTTTAGGAGTACAGATTTGCTGGATTTGATAGGATGAGTGTGTCG 1476
```

QY 301 AenGluileAlaGlyAlaLeuLysBrgSerGlnPheLeuValProMetValSer----- 318  
DB 1477 AATGAGATTGCCGGTCTTGAAGCGGTCAAGATTCTTGTCCCTAGTCTCAGGTACC 1536  
QY 318 ----- 318  
DB 1537 ATATTCTCTCTCACTCGGTGATTTCAATTCGAAAGGTGGTCTCTTTTGTGACATCATC 1596  
QY 318 ----- 318  
DB 1597 GACCAACATCAAGTTCCATCTTTGTTTTCGATAGCTTGATGGTATTAACATGAGAGAGC 1656  
QY 318 ----- 318  
DB 1657 ACATCAATATTATAGAGTCAATGACTGATGAGCCAAATCCTAGCTAGAAATTAATCTG 1716  
QY 318 ----- 318  
DB 1717 GAAAGAACTTGAACCTCTCAACCATAGGTTTGGTACGAAATTTGTTCTGTCTGCAAGCCA 1776  
QY 318 ----- 318  
DB 1777 AATGATAGGCTATTGCCCTGAATAGTAGTGTCTTGTGGTTTCCATATTTGGAAGTTAAA 1836  
QY 318 ----- 318  
DB 1837 TCATATGACTTAGCTGTGGATACATAATTAAGCTTAAGCAATGCCCACTTAAGAAGTGG 1896  
QY 319 -----GlyileValGluSerSerIleLysArgGlyMet 329  
DB 1897 TACTTACAAATATTCTATTGTCATAGGTATAGTTTGAATCAAGTATCAAGCTGGGAATG 1956  
QY 330 HisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349  
DB 1957 CATATTGAAGCTCTTGAGATGTTTATACCTTTTGGCATGGAGGATAAGTTTTCAGTCT 2016  
QY 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369  
DB 2017 CTAGTTCTAACTTCACTCTTAAGATGAGCAAGGAGTCATTGAGAGGCAAAACGGAAA 2076  
QY 370 AlaGlnSerProLeuAlaPheLys----- 377  
DB 2077 GCCAGTCACCGCTGGCATTTGT-ATGAACCCCTCCCTTGCACATTATGTACTTTATGA 2135  
QY 378 -----GluAla 379  
DB 2136 ACTCTTTATCATCATCTGAGTCTGACCATTTGATATATTATTTCTCAACAGCAAGACGG 2195  
QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399  
DB 2196 GGTACAAAGCAGCTAGCTGTGTATCACTATGTCAGTTATGCGAGTATGAGACTCACAAAGTTA 2255  
QY 400 AspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuLys 419  
DB 2256 GATCTCTCGAAAGAACTACAGAGTGGCAGATCAAGAGCAAAATTTGTAGCTTGGAGAAA 2315  
QY 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaAspSerLeuSerLeuMetGlu 439  
DB 2316 GACACTCTTCAGCTGCACAAAGAGATGGAAGAGCAAGATCTCTCAGTTTAAATGAG 2375  
QY 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459  
DB 2376 GAAGCGCACTTGCCAAAGAGATGTATTAACCAACAGATAAAACGTTCCAAAGGTTGTACCC 2435  
QY 460 MetGluMetProValThrSerSerSerSerTyrSerProIleTyrArgAsnArgSerPhe 479  
DB 2436 ATGGAAATGCCACCACTACTCTTCATCGTATTTCTTCATCTACCGGTATAGAGCTTT 2495  
QY 480 ProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGly 499  
DB 2496 CCTAGTCAAGAGACGATGACCAAGATGAATATCAGCTCTTGTGAGTAGTTACCTCGGC 2555

QY 500 ProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeuPro 519  
DB 2556 CGCTCAACATCTTTCTCTCATCGCTCAGAAATCCCGGATATATGGTTCACCTTCCA 2615  
QY 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539  
DB 2616 CATGGTGGGTTTAGGAGAGTGTATATGCATATGAACATCTGGCCCAAAATTCATCTCT 2675  
QY 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559  
DB 2676 CCAAGTCAAGGACATAGACTTCTCATGACAGTACTCTCGTCTTTGGTTCACGACAGAGA 2735  
QY 560 HisProLeuGlnTyrSerProIleHisGlyGlnGlnLeuProTyrGlyIleGln 579  
DB 2736 CATCCACTACAGTACTCTCTCAATTCATGACAAACACACTTACCATATGATATACAA 2795  
QY 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599  
DB 2796 AGGTTTACAGACATTCACCATCTGAAGAAATATTTGGGTTTATCCAAATCAAGGTCT 2855  
QY 600 ProArgSerAsnSerSerLeuAspProLys 609  
DB 2856 CCTCGCAGTAACCTCATCATTAGACCCAAA 2885

## RESULT 10

F6N23  
LOCUS 91849 bp DNA linear PLN 12-NOV-1999  
DEFINITION Arabidopsis thaliana BAC F6N23.  
ACCESSION AF058919

VERSION AF058919.2 GI:6382042

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 91849)  
AUTHORS Geisel, C.  
TITLE The sequence of A. thaliana F6N23  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 91849)  
AUTHORS Washington University Genome Sequencing Center.  
TITLE The A. thaliana Genome Sequencing Project  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 91849)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (10-APR-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 4 (bases 1 to 91849)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (12-NOV-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT On Nov 12, 1999 this sequence version replaced gi:3047100.  
Submitted by:  
Genome Sequencing Center  
Department of Genetics, Washington University,  
St. Louis, MO 63108, USA  
e-mail: twilson@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

#### NEIGHBORING COSMID INFORMATION:

The 5' clone is F5110, 200 bp overlap; 3' clone is F15P23, 900 bp overlap. Actual start of this clone is at base position 104960 of CELF5110; actual end is at 91149 of CELF6N23

The clone sequenced to the left is F5110. The actual start of this clone is at base position 104960 of F5110; actual end is at base position 91149 of F6N23.

#### NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation).

#### FEATURES

##### Location/Qualifiers

```
1. .91849
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /cultivar="Columbia"
  /db_xref="taxon:3702"
  /map="unknown"
  /chromosome="IV"
  /clones="F6N23"
  /complement(187. .907)
  /gene="F6N23.23"
  /complement(join(187. .347,396. .512,598. .907))
  /gene="F6N23.23"
```

#### gene

#### CDS

```
/evidence=not_experimental
/protein_id="AAC13624.2"
/db_xref="GI:6382044"
/translation="MENLSNPDENDDHOSPRSDONDSOAVETPVYSTMSIDSFVYPR
TCSFTTSGFDIDETMSRCSEAPCNWVLTESKSKSLGLEMSQNECLVQELSE
PELTMRERAKLLGDMGSGKGVCTATISNAITNLNDKNDTSFNDVCSYVW
TESVRAVRNEESIEVERNELSFICMLHCRHP"
/complement(4919. .7497)
/gene="F6N23.22"
/complement(join(4919. .4993,5078. .5488,5567. .6055,
6157. .6213,6498. .6594,6935. .7222,7349. .7497))
/gene="F6N23.22"
```

#### gene

#### CDS

```
/evidence=not_experimental
/protein_id="AAC13623.1"
/db_xref="GI:3047112"
/translation="MSLTMDGVAAAGRSKQNSLLRKLALAVRSVQMSYAFWSS
SLTPGVLEWEGCYGDMKKKSYSHYKGLQSKELKLYLMLEDSGTTWT
THNLNDDNCHSTSMWLDLSDEWYVLSMSVFPSPCLPGASATGETIWL
CNAOYAENKLSRSLIARTVVCFLAGVLELGYTELISDHNLRLNKSLMEISAH
QNDDEKKNKIKSEEEKQLPLGLSDEDLHYKRIISVLYNSADRSGKNKRIHRCP
NIVTSEPSSFLRWQCEQQVGFQKKSQNVKLIHDVPLMKRMFPQSGNLN
QDPSDRKENEKESVLTMTPTNEVDKESILNNTIKYQLEEARVEELSCGVS
FVERQRTENLNDLSIETSGNYDDSKIDDSGETQVTPFRDKTHLRVKLKETE
VIVRCSYRQYIVADIMETLSNLHMDAFSVRSHTLNKFLTLNLKAKFRGAASVGM
IKRLRRVIGDLF"
10235. .12675
/gene="F6N23.1"
```

#### gene

#### CDS

```
/join(10235. .10495,10693. .10896,10998. .11192,11341. .11505,
11633. .11893,11998. .12216,12315. .12534,12617. .12675)
/gene="F6N23.1"
/translation="MAIRLNHSVIPSVYKLGAPTRVARSLSLPPSVGDWRGVSTFGA
RPLVLAKVLAESTEDRVPIDDDDSDTQVSVFHLDDSLILGIDMNSVVEPEE
note="similar to the family of glycosyl hydrolases"
```

#### gene

#### CDS

```
/complement(13027. .14791)
/gene="F6N23.21"
/complement(join(13027. .13335,13480. .13615,13712. .13840,
13943. .14001,14087. .14791))
/gene="F6N23.21"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC13622.1"
/db_xref="GI:3047111"
/translation="MSILCCPVLECYVCLGCTHMLMKKCLYSAGHSESNWGLATSE
PEPFRICRLIAYVEENHDPWAPFDGIDENHVLKDYQDTEGRVTPMYILD
HENGDVLAIRGLNAKECAVALLDNKLGQTKFDGGVYVHNLKAAWYFEEHVVL
RELLEANSYSLSITFVHSLGAGVSLVLFVIONRVLRIERKRIRCFATAPRCMS
LHLAVYADVNSVLQDDLPRTTALENVFKSIIICLLCLCTCKDTPFERKL
KDARLYAPGRLIHIVVKRLGRYPVVRATVPDGRFEQIVLSNATADHAIWI
ERESQALDKIVRQKSIIVEDHDEYRAIMKAASINLPMSPPSIGTIFDTEGESSA
GSGMEGSPSGWSFKMRKRKQDFIDCHFPVNDNSEHMIIFKNQESQALLCGOET"
15539. .16687
/gene="F6N23.2"
/join(15539. .15554,15672. .15764,15953. .16026,16133. .16208,
16307. .16353,16466. .16687)
/gene="F6N23.2"
/translation="Similar to protein kinase; coded for by A. thaliana
cDNA H36947; coded for by A. thaliana cDNA H37158"
```

#### gene

#### CDS

```
/codon_start=1
/protein_id="AAC13612.1"
/db_xref="GI:3047101"
/translation="MNTESVVEFLGNVTLORLPSSSLKRISEVVVKYGRDGVVVR
ENQVDGVIFLLQGVRAQVRLASGEENYQEPFLKRYDFGFGIFGDVYSADVAVTE
LTCLLMSDHRALLLEIKVSQSDKERCLVEDILYLEDVDRVSPNCPSPSSISVLMIK
VLTIFYHHSIESEL"
16976. .19151
/gene="F6N23.3"
/join(16976. .17111,17262. .17415,17785. .17889,18316. .18417,
18618. .18666,19056. .19151)
/gene="F6N23.3"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC13635.1"
/db_xref="GI:3047124"
/translation="MEIYSNHCDTTVYASQALAAASKVTETMKIYVNFHCYFLLVGDI
NIPILYDNEURDGNPATRSVDAPQKGTITLTFASQVSCYIPEIHFYSLDVRNQ
VATEISVPPPIDIRFCEPNRSYKONKSPRCVAVASDLIFATISLPHRREGMSVAA
LSLDSHMPFRPRADDMLLFVIVSFATSGRFAIGKFNKKGVRCCIEE"
19151. .21366
/gene="F6N23.20"
/complement(join(20843. .20944,21002. .21056,21103. .21155,
21292. .21366))
/gene="F6N23.20"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC13621.1"
/db_xref="GI:3047110"
/translation="MVATAVSFYLLTGDYGEPPNVLDPLYSSMLILVIVLIRMF
SVWVMLHFLMVIIEHWIKRKLAAQDSVKQIFGPKGSPSGKESDTAK"
```

#### gene

#### CDS

```
/complement(22410. .22521,23220. .23538,23620. .23939,
24009. .24137,24240. .24352,24439. .24548,24639. .24753))
/gene="F6N23.19"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC13637.1"
/db_xref="GI:3047126"
/translation="MSPCSYINVSUKPFFHLMQDQILTNVVKYQGRNWKRTAECILPG"
```

#### gene

#### CDS

```
/complement(20843. .21366)
/gene="F6N23.20"
/complement(join(20843. .20944,21002. .21056,21103. .21155,
21292. .21366))
/gene="F6N23.20"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC13621.1"
/db_xref="GI:3047110"
/translation="MVATAVSFYLLTGDYGEPPNVLDPLYSSMLILVIVLIRMF
SVWVMLHFLMVIIEHWIKRKLAAQDSVKQIFGPKGSPSGKESDTAK"
```

#### gene

#### CDS

```
/complement(22410. .22521,23220. .23538,23620. .23939,
24009. .24137,24240. .24352,24439. .24548,24639. .24753))
/gene="F6N23.19"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC13637.1"
/db_xref="GI:3047126"
/translation="MSPCSYINVSUKPFFHLMQDQILTNVVKYQGRNWKRTAECILPG"
```

SEENNRNDVOCQRHUKVLDP SLOQKAWKBEDELLSELVDYDYNDRPWSKISLEL  
PRTGCRORRHHNNLPILKSPVTRREBLLVQAQNGKQWAEJAKULPQRTENN  
IKNHWGKVRKLEQFPNNLFGVWYSPGSPGPNVFNQRTWATKISQIKGAA  
KSPORDELDTLGNWRSSTSSILRGESBVSVDVCAARLCAKEFTQNSNDVE  
CQXEBRMEKRLRMAARTFTDTS14SKTSPASGLURKQKQYDFTDARSHMSSE  
DHYSVSAVSPBVKVNTCSGSKPLERSLDFTLWDBHGRNGINGVNSVRILPQKS  
DLSKSIYSPBWT, #

[illegible]

US-09-890-475-1 (1-609) x F6N23 (1-91849)

Qy	1	MetSerAsnTyrProProThrValAlaAlaGlnProThrThrThrAlaAsnProLeuLeu	20
Db	56839	ATGTCCAATTATCCACCGAGCGTGGCGGGCAACCCACACACGCGGAATCCACTGCTG	56899
Qy	21	GlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGluThrGluSer	40
Db	56899	CACGGACATCAATCTCTGAAACAGCGACGAAGAATACCGAAGTTGTGAAACAGAGTCT	56959
Qy	41	ThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLysSerIleAsp	60
Db	56959	ACAAGTATGGACATTACGATCGGTCAATTAAGCAGCCTCAATTTTGAATCCATAGAC	57018
Qy	61	GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln	80
Db	57019	GAATTAGCTCGCTTTTCAGTTGCGAGTGGAAAAATTCAAACGCCAATTCGATGATCTTCAG	57078
Qy	81	LysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVal	100
Db	57079	AAGCATCTCAGTCAATCGAAACGCAATTGATTCMAACTCGAGGTAAACGGCGTTGTC	57138
Qy	101	LeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnVal	120
Db	57139	CTCGCCGCGGGAACAATAATTTCCATCAGCCGATGTTATCGCTCCGCGGAACAATGTA	57198
Qy	121	SerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSer	140
Db	57199	TCTGTAAACCAACCGTCACTGTGAGCCAAACCGTCTCAGAGATTGTACCGAGACGTGCG	57258
Qy	141	AsnLysProGluGlyGlyArgMetCysGluMetCysSerLysGlyLeuArgLysTyr	160
Db	57259	AATAAACCGGAGGGGAAACGTAATGTGTGAGTTGATGTGTAGCAAAAGGTTCGCTTAATAC	57318
Qy	161	IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerSerAlaLeuLys	180
Db	57319	ATATACGCGAATATCTCTGATCAAGCTAATGTTAATGGAAGAGATTCTCTCAGCTTTGAAA	57378
Qy	181	LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly	200
Db	57379	TTGGCCACAGAGCAGCGAAGTTTGATTGGATTGATTGGCAAGTTTACTTACCAAGGG	57438
Qy	201	ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle	220
Db	57439	CGTAGACATTTACTAAAGAGTCGCTATGAGCTCTGCGAGACAAGTTTCGTTCTTATA	57498
Qy	221	LeuGluSerPheIleLeuMetProAspArgGlyLysGlyValLysIleGluSerTrp	240
Db	57499	CTGAGTCTTTTCTTCTAAATCGCTCGTGGTAAAGGGAAGGTGAAGATTGAGAGTTGG	57558
Qy	241	IleLysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGly	260
Db	57559	ATTAAAGATGAGCGCGAGACGGCTGCTGTGCTTGGAGGAAGGTTTGATGACTGAGGA	57618
Qy	261	GlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe	280
Db	57619	GGATTAGCTCGGCTGAGAAAATCGATGCAAGGGGTTTCGTTTTTACTAGTTCGCTGTTT	57678

Qy 485 AspaSpGlnAspGluIleSerAlaLeuValSerSerTyrIleuGlyProSerThrSerPhe 504  
Db 58757 GATGACCAAGATGAATATCATGCTCTTGTGAGTAGTATTACCTCGCCCGTCAACATCTTTT 58816  
Qy 505 ProHisArgSerArgArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGly 524  
Db 58817 CCTCATCGCTCAAGAAGATCCCGAATATATGGTTCCTCCATGCTGGGTAGGA 58876  
Qy 525 ArgSerValTyrAlaTyrGluHisLeuAlaProHisSerTyrSerProGlyHisGlyHis 544  
Db 58877 AGAAGTGTATATGATATGAACATCTCGCCCAATCATATTCTCAGATCCAGACAT 58936  
Qy 545 ArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyr 564  
Db 58937 AGATTCATCGACAGTACTCTCGCTCTTGGTTCACGACAGACATCCACTACAGTAC 58996  
Qy 565 SerProProfileHisGlyGlnGlnGlnLeuProTyrGlyTyrGlnArgValTyrArgHis 584  
Db 58997 TCTCTCTCAATTCTGACACAAACAGTTCACCATATGTTATACAAAGGTTTACAGACAT 59056  
Qy 585 SerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerAsnSer 604  
Db 59057 TCACATCTGAAAGAAATATTTGGTTTATCCATCAAAAGTCTCTCCAGTAACTCA 59116  
Qy 605 SerLeuAspProLys 609  
Db 59117 TCATTAGACCCCAAA 59131  
RESULT 11  
ATCHRIV2 197975 bp DNA linear PLN 16-MAR-2000  
LOCUS Arabidopsis thaliana DNA chromosome 4, contig fragment No. 2.  
DEFINITION AL161472  
ACCESSION AL161472  
VERSION AL161472.2 GI:7267383  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 197975)  
Lamar, B., Stoneking, T., Stumpf, J., Mewes, H.W., Lemcke, K. and  
Mayer, K.F.X.  
AUTHORS Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 197975)  
EU Arabidopsis sequencing project.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer  
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:  
lencke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project  
Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge  
Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,  
E-mail: michael.bevan@bbsrc.ac.uk  
COMMENT Information on performance of analysis and a more detailed  
annotation of this entry and other sequences of chromosomes 3, 4  
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>  
this fragment has an overlap with ATCHRIV1 at the 5' end and an  
overlap with ATCHRIV3 at the 3' end.  
FEATURES  
source  
1..197975  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
10352..17547  
/gene="AT4g00450"  
join(10352..10384,10432..10545,10605..10709,10849..11018,  
11079..11423,11475..13674,13735..13823,13921..16530,  
16677..16862,16976..17042,17092..17547)  
/gene="AT4g00450"  
/note="Contains Aminoacyl-transfer RNA synthetases  
class-II signatures AA1343-1352; Prokaryotic membrane  
lipoprotein lipid attachment site AA349-359  
contains Ser 50:AI999580.1"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAB80854.1"  
/db\_xref="GI:7267384"  
/db\_xref="SPTREMBL:O23057"  
/translation="MLSPLFVISIGAVKVKLRAINESRALRKVKIKRLSDSNPQFIF  
CAEPKAGVYGVPLSGSLCKPPEORSCEETKRWIESLSQHKELSLANIPG  
YRKTLFVLRNNVPLLRATWFIKTYILNQHVCNAINWCDLVLLFLEHLEHKKLL  
GVNGKMLNLCNTSNWFCGIAHFLSKLEIGHERCFIQDCKRFLVQHQAFAR  
KHLIYNGIWCVFYSGMGLFFLLISLIGSSSVFVNFHALLQKEIFELQL  
LLPIYGVLESIVLSQTVSLAVAFIOEPAPGGSDVDNRRATYLLALLLEMR  
VLVAAPDTFVADFFPLPSPVAACPNDSYTKAVENKLSNAGAEIQAQOGRG  
VLSRPEFLSPDYTTIORSADDLAKIAGYPOHNAKAVQALDKALSDGIDRAAYS  
VLFEDLNGAVDEAWITDVSCLRSSLRWIGAI STFCVCSVEFLIENATCDPFRFAG  
VPKIKFSGRKDCQVILVQLKQLIGESTARKKGNCKNNELGVSKPSGSDAFE  
SPGPHLDIIVICDOHEVHGKAKRLQLLFEELRSGIFNFPIATVRLIVSGMDVIG  
PAVDPERMRHRLIKQLPGCFVHETLEEAQLFGDKLSEAVRTYSNRRLLRELLV  
EKGYNNLVLSQSKKISTLSLVIPFRACNAKSNKSPKRTKSSVDIRELKERI  
SALLQFPQMSCGVETPVRFQNSVGRSSGVSKMDQPEATPGCEDCRARPRKMD  
EKSCYQCNSTPIASDEEDNMMIKKSKTVESLSKVDPOIETKQVPRGRQWARKTQ  
LAQLAARIESGQASTSHVCDNKVSCPHGPGVGEKGVVDVFRISTPVDYMSVGN  
SLKQLQFVKRSIAWLTAVRLVLEPKSSVAVGPNCGAPVEEKNTIEMKLGAD  
LSRATKIATAVSALVIGSIHGVITLERIVGLLKLKDYLDVQVFRVTKSSNSAS  
MSKSPESPIEVYHNFRLVGNCKTVSEGLVLELVGESSVAVSGRMGLPLKLVFP  
PAYSIATVLRFPVFNSSNSNSVHEDTHRLYQSLTMAFHDVVKHLPDRCVCFRTQ  
LYELIADSTDAEFASVESHGLMELKSAFAPLARLFLNLDICKVPSGSHGEG  
VSEAKNRHQNGTKLVKLVSLDCLQPAKEHMQWVELRLINLEQALAEKLENDHML  
TDAIRSCPTSEKPDASEKNFIQLLTRLVDPAPVLFSEVHLFGRSVEDSMLK  
QAEVLAGDVLFGKRTIROLIIVGESKGLPTKPOFWKPMGWCNSSSSSDHTANKAG  
KXKPEITSIEGEVIEGSGSRKVLPRVLDENSPVGYGITTERAVQLVLPIDQ  
SDDSEITSFVNEVLVRFNIEQQSSVTRNTSNKMGTSAGEISESSKSGTKRGL  
RGSPSLARRSANTDTSPFPAALRASMLRLQLFLLRPVLCGSPFNKTHAL  
ASTVLRIGSRVYEDYAVCSFRSELKASTETSDPSSMADLSSEVLFDRRLTGLS  
TNFKSLKLVCLLGCKNELSRVLPDTRIRIQAAAPILLPSRLCSLQCPHSVPPTAL  
TLVQPSGTAAGTANQRNSPAISKSGTAAAGKLPMTLLAPHQOQEAONTVDVDPWT  
LEDGSSGLSSSNASNSDMMANLRATCWLKAVRVRTDITLVVSVDSDS"  
10352..10384  
/gene="AT4g00450"  
/number=1  
10385..10431  
/gene="AT4g00450"  
/number=1  
10432..10545  
/gene="AT4g00450"  
/number=2  
10546..10604  
/gene="AT4g00450"  
/number=2  
10605..10709  
/gene="AT4g00450"  
/number=3  
10710..10848  
/gene="AT4g00450"  
/number=3  
10849..11018  
/gene="AT4g00450"  
/number=4  
11019..11078  
/gene="AT4g00450"  
/number=4  
11079..11423  
/gene="AT4g00450"  
/number=5  
11424..11474  
/gene="AT4g00450"











US-09-890-475-1 (1-609) x AY092693 (1-558)

QY 34 LysIleValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnPro 53  
 Db 1 AAGATTGTCGAAACAGAGCTTACAGATGATGACATTCAGATCGGTCAATCTAAGCAGCCT 60  
 QY 54 GlnPheLeuLysSerIleAspGluLeuAlaIaPheSerValAlaValGluThrPheLys 73  
 Db 61 CAATTTTGAATTCCTAGAGCAATTAGTCGCTTTCAGTTGAGTGGAAACATTCCAAA 120  
 QY 74 ArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLys 93  
 Db 121 CGCAATTTCGATGATATTCAGAACGACATCGAGTCAATCGAAACGCAATTGATTCCAAA 180  
 QY 94 LeuGluSerAsnGlyValValIleAlaIaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113  
 Db 181 CTCGAGATACGGGGTGTCTCTCCCGCGGGAACATAATTTCCATCAGCCGATGTTA 240  
 QY 114 SerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133  
 Db 241 TCGCCTCCGCGAACAATGTATCTGTAGAAACCAACCGTCACTGTGAGCCACCGTCTCAG 300  
 QY 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153  
 Db 301 GAGATTGTACCGGAGACGTCAATAAACCGAGGGGGACGTATGTGTGAGTTGATGTGT 360  
 QY 154 SerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173  
 Db 361 AGCAAGGTCTCGGTAAATACATATACCGGAATATCTCTGATCAAGCTAAGTTAATGGAA 420  
 QY 174 GluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIle 193  
 Db 421 GAGATTCCTTCAGCTTTGAAATTGGCCCAAGGAGCCAGCAAGTTTGTATTGCTGATTATT 480  
 QY 194 GlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAla 213  
 Db 481 GGCAAGTTTACTTACAGGGCGGTAGAGCATTTACTAAGAGTCGCCTATCAGCTCTCGG 540  
 QY 214 ArgGlnValSerLeuLeu 219  
 Db 541 AGACAAAGTTTCGCTTCTT 558

Search completed: February 28, 2004, 04:51:32  
 Job time : 5820 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 27, 2004, 21:18:27 ; Search time 574 Seconds  
(without alignments)  
4507.235 Million cell updates/sec

Title: US-09-890-475-1

Perfect score: 3104

Sequence: 1 MSNPPTVAAQPTTIANPLL.....RYGLSNQRSPRNSSLDPK 609

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n\_model -DEV=xlp  
-Q/Cgn2\_1/USPTO\_spool\_p/US09890475/runat\_24022004\_135311\_5281/app\_query.fasta\_1.775  
-DB=N\_Geneseq\_25Jan04 -QFMI=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCUT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09890475 @CGN 1 1.885 @runat\_24022004\_135311\_5281 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3104	100.0	2257	3 AAA63669	AAA63669 cDNA sequ
2	2981	96.0	3761	3 AAA63668	AAA63668 H51 genom
3	349	11.2	2155	3 AAC33344	AAC33344 Arabidops
4	340	11.0	2158	3 AAC46135	AAC46135 Arabidops
5	337	10.9	1632	9 ADC03365	ADC03365 Rice flow
6	329	10.6	1802	9 ADC03469	ADC03469 Maize flo
7	296.5	9.6	1964	3 AAC49876	AAC49876 Arabidops
8	296.5	9.6	1970	3 AAC35353	AAC35353 Arabidops

9	293	9.4	1377	9 ADC03367	ADC03367 Rice flow
10	288.5	9.3	1865	3 AAC38630	AAC38630 Arabidops
11	234.5	7.6	1524	6 AB213112	AB213112 Arabidops
12	224	7.2	947	9 ADC03447	ADC03447 Wheat flo
13	152	4.9	7785	4 ABL17653	ABL17653 Drosophil
14	151	4.9	3312	2 AA232435	AA232435 Rat TAO1
15	151	4.9	3312	2 AA232435	AA232435 Rat TAO1
16	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
17	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
18	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
19	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
20	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
21	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
22	148.5	4.8	3048	8 AAU62252	AAU62252 Human TAO
23	145.5	4.7	2760	6 ABK32901	ABK32901 DNA encod
24	145	4.7	2760	6 ABK32901	ABK32901 DNA encod
25	145	4.7	2760	6 ABK32901	ABK32901 DNA encod
26	144.5	4.7	3003	8 AAL62251	AAL62251 Human TAO
27	144	4.6	2684	6 ABSS1734	ABSS1734 Human nov
28	144	4.6	2708	9 ADA08249	ADA08249 DNA encod
29	144	4.6	2756	6 ABK70288	ABK70288 Human lun
30	144	4.6	2756	6 ABK70288	ABK70288 Human lun
31	143	4.6	3270	7 ABT16471	ABT16471 Human int
32	143	4.6	18506	4 ABL05592	ABL05592 Drosophil
33	142	4.6	3114	2 AA02800	AA02800 Human RHA
34	142	4.6	3114	2 AA02800	AA02800 Human RHA
35	141.5	4.6	4772	4 ABL03209	ABL03209 Drosophil
36	141	4.5	2020	6 ABSS1735	ABSS1735 Human nov
37	140.5	4.5	2511	4 AAK51630	AAK51630 Human pol
38	140.5	4.5	2609	9 ADA08250	ADA08250 DNA encod
39	140.5	4.5	3587	9 ADB63157	ADB63157 Human cDN
40	140	4.5	3718	6 ABK63564	ABK63564 Rat sequ
41	137.5	4.4	6558	3 AAZ45602	AAZ45602 cDNA sequ
42	136	4.4	6455	9 ACF79923	ACF79923 Breast ca
43	136	4.4	6455	9 ACF79923	ACF79923 Breast ca
44	135	4.4	7080	5 AAS85528	AAS85528 DNA encod
45	135.5	4.4	14278	4 ABL06088	ABL06088 Drosophil

#### ALIGNMENTS

RESULT 1  
AAA63669  
ID AAA63669 standard; cDNA; 2257 BP.

XX AC AAA63669;

XX DT 04-DEC-2000 (first entry)

XX DE cDNA sequence of the FRI gene of Arabidopsis.

XX KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;  
XX flower initiation; stem elongation; flower production; ss.

XX OS Arabidopsis sp.

XX FH Key Location/Qualifiers

XX FT CDS 362..2191

XX FT /\*tag= a

XX FT /product= "polypeptide which alters flowering time"

XX PN WO200046358-A2.

XX PD 10-AUG-2000.

XX PF 25-JAN-2000; 2000WO-GB000197.

XX PR 05-FEB-1999; 99GB-00002660.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Johanson U, West J, Dean C;

XX



```

QY 581 ValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSerPro 600
AAA63668
DB 2102 GTTTACAGACATTCCACCATCTGAAGAAAGATATTGGTTTATCCAAATCAAAGGTCTCCT 2161
QY 601 ArgSerAsnSerSerLeuAspProIys 609
DB 2162 CGCAGTAACATCATATTAGACCCCAAA 2188
RESULT 2
ID AAA63668 standard; DNA; 3761 BP.
AC AAA63668;
DT 04-DEC-2000 (first entry)
DE H51 genomic sequence of first 17 kb of 84M13 encompassing FRI gene.
KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; ss.
XX Arabidopsis sp.
XX WO200046358-A2.
XX 10-AUG-2000.
XX 25-JAN-2000; 200OWO-GB000197.
XX 05-FEB-1999; 99GB-00002660.
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX Johanson U, West J, Dean C;
XX WPI; 2000-532899/48.
XX New nucleic acid derived from the FRI locus of a plant, e.g. Arabidopsis,
PT encoding a polypeptide capable of specifically altering the flowering
PT time of a plant.
XX Claim 4; Fig 4; 73pp; English.
XX The present sequence represents the (late flowering) H51 genomic sequence
CC of the first 17 kb of cosmid 84M13 encompassing FRI (one locus-FRIGIDA)
CC locus of Arabidopsis. The FRI gene encodes a polypeptide capable of
CC specifically altering the flowering time of a plant. The FRI
CC polynucleotide is used to transform plants, so that the flowering time of
CC a plant is altered. This is used, for example, for plants in which the
CC leaves or tubers are a commercial product, where it is desirable to avoid
CC 'blotting' (initiation of flowers and stem elongation) at too early a
CC stage. Conversely, it may be desirable to alter flowering under certain
CC circumstances e.g. to vary flower production across the seasons
XX SQ Sequence 3761 BP; 1174 A; 717 C; 775 G; 1095 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6.4e-232 Length: 3761
Score: 2981.00 Matches: 608
Percent Similarity: 78.96% Conservative: 0
Best Local Similarity: 78.96% Mismatches: 1
Query Match: 96.04% Indels: 162
DB: 3 Gaps: 2
US-09-890-475-1 (1-609) x AAA63668 (1-3761)
QY 1 MetSerAsnTyrProProThrValAlaAlaGlnProThrThrAlaAsnProLeuLeu 20
DB 574 ATGTCCAAATTATCCACCGAGGTGGCGGCAACCCCAACAGCGGCAATCCACTGCTG 633
QY 21 GlnArgHisGlnSerGluGluArgGluLeuProIysIleValGluThrGluSer 40

```

```
QY 318 ----- 318
DB 1774 AATGATAGGCTATTGCGCTTGAATAGTGTCTTGTGGTTTCCAAATATTGGAAGTTAA 1833
QY 318 ----- 318
DB 1834 TCGTATGACTTAGCTGTGGATCTAAATTAAGCTTAAGCAATGCCAACTCTAAGAACTGG 1893
QY 319 ----- 319
DB 1894 TACTTACACAATATCTATTGCTCATAGTATAGTGAATCAAGTATCAAGCGTGGATG 1953
QY 330 HisLeuGluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAla 349
DB 1954 CATATTGAAGCTCTTGAGATGTTTATACCTTTGGCATGGAGGATAAGTTTTCAGCTGCT 2013
QY 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLys 369
DB 2014 CTAGTTCTAACTTCTTAAAGNAGAGCGAGGATCATTTGAGAGGCGCAAAACGGAA 2073
QY 370 AlaGlnSerProLeuAlaPheLys ----- 377
DB 2074 GCCCAGTCACCGCTGCCATTCTGT-ATGAACCCCTTCCTTGCACATTATGTACCTTTTATGA 2132
QY 378 ----- GluAla 379
DB 2133 ACTCTTTATCATCTGAGCTGACCATGATGATATATTTATTTCTCAACAGAAAGACGG 2192
QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399
DB 2193 GCTACAAAGCAGCTAGCTGTGTATCATCAGTATGAGTATGAGGACATCACAAGTTA 2252
QY 400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLys 419
DB 2253 GATCTCGGAAGAAGACTACAGATGGCAGATCAAGAGCAAAATTTGTAGTTTGGAGAA 2312
QY 420 AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGlu 439
DB 2313 GACACTCTTCAGCTCGCAAGAGATGAGAGAGAAAGCAAGATCTCTCAGTTTAAATGGAG 2372
QY 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459
DB 2373 GAAGCGCGCACTTGCCAAAGAGATGATATACCAACAGATAAACGTCACAGTTGTACACC 2432
QY 460 MetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSerPhe 479
DB 2433 ATGGAAATGCCACAGTAATCTTCTCATCGTATTTCTCTATCTACCGTGATAGAGCTTT 2492
QY 480 ProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGly 499
DB 2493 CCTAGTCAAAAGAGACGATGACCAAGATGAAATATCAGCTCTTGTGAGTAGTTACCTCGGC 2552
QY 500 ProSerThrSerPheProHisArgSerArgArgSerProGluTyrMetValProLeuPro 519
DB 2553 CCGTCAACATCTTTTCTCATCGCTCAAGAGATCCCGGAATATATGTTCCACTTCCA 2612
QY 520 HisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSer 539
DB 2613 CATGGTGGGTAGGAAGAAGTGTATATGATATGAACATCTGGCCCCCAAAATTCATCTCT 2672
QY 540 ProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArg 559
DB 2673 CCAGGTCCAGGACATAGATCTCATCGACAGTACTCTCCGCTCTTGGTTCACGACAGAGA 2732
QY 560 HisProLeuGlnTyrSerProProIleHisGlyGlnGlnLeuProTyrGlyIleGln 579
DB 2733 CATCCACTACAGTACTCTCTCCAAATTCATGACACACAGTATCATATGTTATACAA 2792
QY 580 ArgValTyrArgHisSerProSerGluGluArgTyrLeuGlyLeuSerAsnGlnArgSer 599
DB 2793 AGGGTTTACAGACATTCACCATCTGAAGAAGATATTTGGGTTTATCCCAATCAAAGGTCT 2852
```

600 ProArgSerAsnSerSerLeuAspProLys 609  
2853 CCTCGCAGTAACCTCATCTAGACCCCAA 2882

RESULT 3  
AAC33344  
ID AAC33344 standard; DNA; 2155 BP.  
XX AAC33344;  
AC AAC33344;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 2690.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 03-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.



Db	392	GCATTTTCATGGCGCTTGAGAGGCTCTTTGAAGAGAGAGGTTTCATGAGCTTGAGATCAAGA	451
Qy	80	nLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValVa	100
Db	452	AAAGAGTAGTGAACCAACCAAGAAAGCTCAAGAACTGTTGGAGAAAAAAGGCGAGC	511
Qy	100	lLeuAlaAlaArgAsn-----	105
Db	512	TGTTGAGGCTTAAGAGAGAGCGCGCTTGAAAGCGCTTCAGAAAGAGAGACCGACGTAT	571
Qy	106	-----AsnAsnPheHisGlnProMetLeuSerPr	115
Db	572	GTTTACTATAAACAGTCTTTGGATAAGTATAACAAC-----GCTCCTGTCCAGTAAACC	625
Qy	115	o-----ProArgAsn-----AsnValSerValGl	123
Db	626	TTCTGTTGGTGAACGTTGGCCACAAAATGCTGTAGAGAGTATCAATGCTTTCTGCTGC	685
Qy	123	uThrThrValThrValSerGlnPro-----SerGlnGluIleValPro--	137
Db	686	TGATAGTATTACCGATGATTAATCCGATGTTATGTGCAAGATGTTTCAGATCTTCACCTGT	745
Qy	138	-----GluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCy	153
Db	746	GATCGGAAATATGACGTGAAGCTTATCCACAGTTACTCAAACTGTGCGGGATATGGA	805
Qy	153	sSerLysGlyLeuArgLysIleTyAlaAsnIleSerAspGlnAlaLysLeuMetGl	173
Db	806	CTCAACGGGTCTCCATAAGTTTGTATCAGATACCGCAAGACCTTGATCATCTCACTAAGGA	865
Qy	173	uGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIl	193
Db	866	GGAGATTCCTATGCGTTCAGGGCTGCAGCAACCCAGCACCTTAGTGTGACATCTTT	925
Qy	193	eGlyLysPheTyr-----LeuGlnGlyArgArgAlaPheThrLysGl	207
Db	926	GGAAAGGTTTACCTATGAGGACCAACTGCTGATGGAAAA-----AAAGA	973
Qy	207	uSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGlu-----SerPheLe	225
Db	974	CGCTAACCTCTGGCGATGGCAGAACCTGTATCATGTTGATGGATGCCCTTAGCATACT	1033
Qy	225	uLeuMetPro---AspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGl	244
Db	1034	GTTATCTCGTGTGACCGCAACTCCCTTGCTGTTGTTCTCTCAAAAATGTTTAAGCATAG	1093
Qy	244	uAlaGluThrAlaAlaValAlaIleArgLysArgLeuMetThr---GluGlyGlyLeuAl	263
Db	1094	AGCCAAACTATTCCAGAGATGGATCCACTCTGGAAATCTCTGACATGATGCTGTTG	1153
Qy	263	aAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPheGlyValPr	283
Db	1154	CAATGGTAATTCCTTTGGAGGCTCATGCATCTCTCAACTGCTGCCACTTTTGTATTGT	1213
Qy	283	oSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIl	303
Db	1214	TGCTGATTTTAAAGAGATGAATCACTTAAAGCTGATCCCATGGTTTCAGCTGCCGCTCA	1273
Qy	303	eAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSe	323
Db	1274	GGCAGCTGAGCTCTGCCGTTTCATCTGGATTAGCTGAGAAAAAATGCTGCTGTTGAAGT	1333
Qy	323	rSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGl	343
Db	1334	TCTCGTGAACAGTGGAAACAGATTGATGGGTTTAACCTGGCAATTCGTTTGAACCTCAC	1393
Qy	343	uAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPh	363
Db	1394	AGAACGTTCTCACCTCTCATTTACTAGTAATCTTACTTGATAGAGCAAGCGCATCTTC	1453
Qy	363	eGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGl	383
Db	1454	CCCCCAAGGACAGCTCGCAATGATCATCTCTGCTGTTTTCAGATGAGTTCAATGAGCGGA	1513

Qy	393	nLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLys	403
Db	1514	GCATTATAGGCCTCAAAACCGGTATAAGTGTATCGAAGAGCATAGCGTGGAA	1567
Qy	403	sGluLeuProGlyTTrpGlnileLysGluGlnileValSerLeuGluLysAspThrLeuGlu	423
Db	1568	GCAGTACCCA	1606
Qy	423	nLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluLalaLalaLe	443
Db	1607	GCCTCGAAGAGCCAAAGCGGACAAAAGAGAGCTACAGAACCCATGAA	1655
Qy	443	uAlaLysArgMetTyrAsnGlnGlnileLysArgProArgLeuSerProMetGluMetPr	463
Db	1656	-----CCTCAGCCAAACACACCGTGGTGCT-----CAACCC	1687
Qy	463	oProValThrSerSerSer-----TyrSerProIleLys	474
Db	1688	CCGAGTCACTGACCAACCAACATCAACACACAGCAGGATACCGTACAGTCAT	1747
Qy	474	rArgAspArgSerPheProSerGlnArgAspAspGlnAspGluileSerAlaLeuVa	494
Db	1748	CCCTGAAAGG--TATCCGAGATGCTATGACACAGACCGTTCCTTAGCGGTCCAAT	1804
Qy	494	lSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgArgSerProGluTyr	514
Db	1805	CATGGCAGCACCAACCTCCACACCTCCTCCTCCTCAGACTTACAGGTTCATCCT	1859
Qy	514	rMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAl	534
Db	1860	-----GCTCCT--GCACATGGAACTTCTACGCGAACTGCTCAGTACCAAGGCTCCTCC	1912
Qy	534	aProAsnSerTyr	538
Db	1913	TCCTCCTCCTTAC	1925
RESULT	4		
ID	AAC46135	standard; DNA; 2158 BP.	
XX	AC	AAC46135;	
XX	AC		
DT	18-OCT-2000	(first entry)	
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 49031.	
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	XX		
PD	06-SEP-2000.		
XX	XX		
PF	25-FEB-2000; 2000EP-00301439.		
XX	XX		
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	23-MAR-1999; 99US-0126284P.		
PR	01-APR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		



PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139452P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147935P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151068P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157533P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	14-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.





```
Db 1390 COTCGTGCACGGATCAGTATATCGCCACATATCTAGC----- 1431
Qy 474 TyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeu 493
Db 1432 TTTTCTGACAAAGTTTC----- 1449
Qy 494 ValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSerProGlu 513
Db 1449 ----- 1449
Qy 514 TyrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGluHisLeu 533
Db 1450 -----TACAGCGCGCT 1461
Qy 534 AlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSer 553
Db 1462 GCGCCACAGAGG-----CATTCCTACCCCTTACGAGAGCGAGTAT----- 1500
Qy 554 LeuValHisGly-----GlnArgHisProLeuGlnTyrSerProProIleHisGly 570
Db 1501 ---GTATATGGTGTGAGGCCCATCATCCCAACAATGATAAGC----- 1542
Qy 571 GlnGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHisSerProSerGluArg 590
Db 1543 ---TCAGCTCCCTATGGCATG-----TCACCTGCCCACACACACA 1578
Qy 591 TyrLeuGly 593
Db 1579 TACTATGGT 1587
```

## RESULT 6

ADC03469

ID ADC03469 standard; DNA; 1602 BP.

XX ADC03469;

DT 18-DEC-2003 (first entry)

DE Maize flowering time-related DNA sequence #4.

KW flowering time; maize; flower architecture; plant biotechnology;

KW commercial plant farming; agriculture; flowering-related protein; ds.

XX Zea mays.

XX WO2003000904-A2.

XX 03-JAN-2003.

XX 24-JUN-2002; 2002WO-EP006969.

XX 22-JUN-2001; 2001US-0300112P.

XX 26-SEP-2001; 2001US-0325277P.

XX 30-NOV-2001; 2001US-0334984P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;

XX Glazebrook J, Katagiri F, Kreis J, Provart N, Ricke D, Zhu T;

XX WPI; 2003-229340/22.

XX New isolated nucleic acid molecule encoding a polypeptide modulating

XX flower architecture and flowering time, useful in the area of plant

XX biotechnology, and commercial plant farming and agriculture.

XX Disclosure; SEQ ID NO 109; 323pp; English.

XX The invention comprises the amino acid and coding sequences of proteins

XX involved in the control of flowering time in rice. The DNA and protein

XX sequences of the invention are useful for modulating flower architecture

XX and flowering time, the DNA and protein sequences are useful in the area

CC of plant biotechnology, commercial plant farming and agriculture. The  
CC present maize DNA sequence shows homology to the rice sequences of the  
CC invention.

SQ Sequence 1602 BP; 442 A; 336 C; 404 G; 420 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	5,08e-17	Length:	1602
Score:	329.00	Matches:	142
Percent Similarity:	40.13%	Conservative:	104
Best Local Similarity:	23.16%	Mismatches:	195
Query Match:	10.60%	Indels:	172
DB:	9	Gaps:	26

US-09-890-475-1 (1-609) x ADC03469 (1-1602)

Qy 36 ValGluThrGluSerThrSerMetPheIleThrIleGlyGlnSerIysGlnProGlnPhe 55

Db 10 ATGGAGTCTGTGCGGCTCTTATGGACTCGACA-----AGCTCCAAGATACAACAGCTT 63

Qy 56 LeuLysIleAspGluLeuAlaAla----- 64

Db 64 CAGAGGCATTCGCTGAGCTTGAGCCAGAGCGCTTCTGCATGAACCTCAAGTGAAG 123

Qy 65 -----PheSerValAlaValGluThrPheLysArgGlnPheAspLeu 79

Db 124 CAGCTTGAAGACCACCTTTCGTGCTTGAACAATCCTCAAGAAAAAGTTTGACGAGCTA 183

Qy 80 Gln----- 80

Db 184 AAAGACGAGAGAGAGAGTTCAGGAGACTGCGCAAAATCAGACAGATCTGGAGCAG 243

Qy 81 -----LysHisIleGluSerIleGluAsnAlaIleAspSerLys 93

Db 244 CGGAGGCGAGCTCTTGTGGCCAGGAACAGACCTCTTTGGAGAGGTTCAGGAGAGAG 303

Qy 94 LeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeu 113

Db 304 GAAGCTGCATTAGCTATGATCTTTGCAAGTCCAGGCTTCTTACCTGCTTCCATC 363

Qy 114 SerProProArg-----AsnAsnValSerValGluThrValThrValSerGln 130

Db 364 AACCCATGAACAAGGACCTGGATAATCTTGGTGTCAAGTGGGCTAAACCGGCTTCTGAG 423

Qy 131 ProSerGlnGluIle---ValProGluThrSerAsnLysPro-----GluGlyGlyArg 147

Db 424 GAGAGTGTGCACCTTGAAGTTGATAATGCTGCAGTGAGGCGCTGTTCTGAGCTTGTGCA 483

Qy 148 MetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAsp 167

Db 484 CTCTGTGAGGAATGAATGAGACGGCTTCATAAGTTTCATATCCGACACAGGAAGAAC 543

Qy 168 GlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys 197

Db 544 TTGCGACCATCCGCGAGGAAATTCCTAGTGCACTGAAGAAAACATCTCATCTCTTATGGC 603

Qy 188 PheValLeuAspCysIleGlyLysPheTyr-----LeuGlnGlyArg 201

Db 604 CTGCTGTGTGATTCCTCTGGAGACTTTTATCTGGAGATAATCTGTATAGATGAATAA 663

Qy 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221

Db 664 -----AAGGATGGTGACCTTCTGGGTGTTAGAGGACATGTTTGTATGTTGATG 711

Qy 222 GluSerPheLeuLeuMetProAspArgGlyLysGlyVal----- 235

Db 712 GAGTCACCTT-----GGACACTGCACCTCTAATGATATACT 747

Qy 236 -----LysIleGluSerTrpIleLysAspGluAlaGluThrAla 248

Db 748 TGTCTCTCTTTGGAGAGGACATGCTTACAACCTAATATCATAGAGCGAGCAAGAGATT 807

Qy 249 AlaValAlaIleTrpArgLysArgLeu-----MetThrGluGlyGlyLeu 262

Db 808 GCATTTAAAGTGAAGTCAAGTTAGATAATCTTGACATTCATGCTAGCAATGGAACTGT 867  
Qy 263 AlaalaalaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyVal 282  
Db 868 -----CTTGAAGCTCAGCGTTCTTCACTCTCTGGCAACTTTGGTATT 912  
Qy 283 ProSerAsnPheArgSerThrAspLeuLeuAspLeuLeu-----ArgMetSer 298  
Db 913 TCTGCTGAATTCATGAAGATGACTTGTGCAAAATGCTTCATATGTCAGTCGTCGCT 972  
Qy 299 GlySerAsnGluLeuAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318  
Db 973 CAGACACTGAATCTTGTGATTCTTGGTTGTGGCAG-----AAATGCCA 1020  
Qy 319 GlyLeuValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyr 338  
Db 1021 GGTGTCTATTGAAGTTCTGTGGAAGTGAAGACCTATTGATGCAATTAATTTGGCTTAT 1080  
Qy 339 ThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMet 358  
Db 1081 GTGTTTGGCTTATGCAACAATTTGAACCACTACACCTTCTTAAGCATATCTAAGGAT 1140  
Qy 359 SerLysGluSerPheGluAlaLysArgLysAlaGlnSerProLeuAlaPheLysGlu 378  
Db 1141 GTTAAAGAA---ANGTCACAGCCAGGAATGTCAAAACCTTCTCTGGAGCACAGAATGAG 1197  
Qy 379 AlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLys 398  
Db 1198 ATGAATGAACGTGAGTTATCTGCGCTGAAATCTGTCATTAAAGTCATTGAAGACACAAA 1257  
Qy 399 LeuAspProAlaLysGluLeuProGlyTrpGlnLysGluGlnIleValSerLeuGlu 418  
Db 1258 CTTGAG-----CAGCAGTACCTT-----GTGGATCCACITTCAG 1290  
Qy 419 LysAspThrLeuGlnLeuAspLysGluMetGluLysAlaArgSerLeuSerLeuMet 438  
Db 1291 AAAAGGGTTTTGCAGTAGAGAAAGCAAGCAGCAAGAGGATGGCTGT----- 1341  
Qy 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSer 458  
Db 1342 ---GAGCAGCA-----AAGCCGCGATGCCAAGAGGCTCTGTGCCAAT 1380  
Qy 459 ProMetGluMetProProValThrSerSerTyrSerProIleTyrArgAspArgSer 478  
Db 1381 GGATCGGCCTTT-----GCAACCGTGCTACAGGCTTGTGTCACAGAGC 1425  
Qy 479 PheProSerGlnArgAspAspGlnAspGluLeuSerAlaLeuValSerTyrLeu 498  
Db 1426 TTCCTACTCA----- 1434  
Qy 499 GlyProSerThrSerPheProHisArgSerArgSerProGluTyrMetValProLeu 518  
Db 1435 -----GCAACCCAGAGAGGATCCATTCAATCTTCCGAGAGG----- 1473  
Qy 519 ProHisGlyGlyLeuGlyArgSerValTyrAlaTyrGlu-----HisLeuAlaPro----- 535  
Db 1474 -----CAGTTGTGTATGCGCTGAGGTTCACTCTCCCTCCCATGATG 1515  
Qy 536 -----AsnSerTyrSer-----ProGlyHisGlyHisArgLeuHisArgGlnTyrSer 551  
Db 1516 ACCTCAGGCTTTACACAGTCAACCTGCCATGGACCT----- 1554  
Qy 552 ProSerLeuValHisGlyGlnArgHisProLeuGlnTyr 564  
Db 1555 -----TATTACGGTAATGGCTACCCAGTTCAGTAC 1584  
RESULT 7  
AAC49876  
ID AAC49876 standard; DNA; 1964 BP.  
XX  
AC AAC49876;  
XX

DT 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 62772.  
DE Hybridisation assay; genetic mapping; gene expression control;  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW Promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0128785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132456P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0138119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.





PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134222P.	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145919P.
PR	18-MAY-1999;	99US-0134768P.	PR	28-JUL-1999;	99US-0145951P.
PR	19-MAY-1999;	99US-0134941P.	PR	02-AUG-1999;	99US-0146386P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146389P.
PR	24-MAY-1999;	99US-0135629P.	PR	03-AUG-1999;	99US-0147038P.
PR	25-MAY-1999;	99US-0136021P.	PR	04-AUG-1999;	99US-0147204P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147302P.
PR	28-MAY-1999;	99US-0136782P.	PR	05-AUG-1999;	99US-0147192P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147260P.
PR	03-JUN-1999;	99US-0137528P.	PR	06-AUG-1999;	99US-0147303P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147416P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147493P.
PR	08-JUN-1999;	99US-0138094P.	PR	08-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138540P.	PR	08-AUG-1999;	99US-0148171P.
PR	10-JUN-1999;	99US-0138847P.	PR	11-AUG-1999;	99US-0148319P.
PR	14-JUN-1999;	99US-0139119P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148565P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148684P.
PR	17-JUN-1999;	99US-0139492P.	PR	15-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139458P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139460P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139461P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139462P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139463P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139458P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139459P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139460P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139461P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139462P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139463P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139463P.	PR	27-AUG-1999;	99US-0151066P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151080P.
PR	18-JUN-1999;	99US-0139763P.	PR	30-AUG-1999;	99US-0151303P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151438P.
PR	22-JUN-1999;	99US-0139899P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140353P.	PR	07-SEP-1999;	99US-0152363P.
PR	23-JUN-1999;	99US-0140354P.	PR	10-SEP-1999;	99US-0153070P.
PR	24-JUN-1999;	99US-0140699P.	PR	13-SEP-1999;	99US-0153758P.
PR	28-JUN-1999;	99US-0140823P.	PR	15-SEP-1999;	99US-0154018P.
PR	29-JUN-1999;	99US-0140991P.	PR	16-SEP-1999;	99US-0154039P.
PR	30-JUN-1999;	99US-0141287P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0141842P.	PR	23-SEP-1999;	99US-0155139P.
PR	01-JUL-1999;	99US-0142154P.	PR	23-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142053P.	PR	24-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142977P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143542P.	PR	05-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0143624P.	PR	06-OCT-1999;	99US-0157865P.
PR	15-JUL-1999;	99US-0144005P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144085P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144323P.	PR	12-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145089P.			



```
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2,94e-14
Score: 296.50
Percent Similarity: 41.10%
Best Local Similarity: 23.13%
Query Match: 9.55%
DB: 3
Gaps: 19

US-09-890-475-1 (1-609) x AAC35353 (1-1970)

QY 59 IleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGln----- 75
Db 131 TTAAGTGTAGT-----CACCGAGTTTCGAGGATTTTCAGAAACAGAGCGTGGT 184
QY 76 -----PheAspLeuGlnLysHisIleGluSerIleGluAsn 88
Db 185 ATGACTAGTGTACTCTCTATGCGCAAGAGCTCTCCGATCACTTCTCTAGAGCAG 244
QY 89 AlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPhe 108
Db 245 AATCTGATGAAGAAATCTGAGCGCTTGAACGAGATGATTGAACTCTAGATAACCGACT 304
QY 109 HisGlnProMetLeuSerProProArgAsnAsnValSerValGluThrValThrVal 128
Db 305 CAGACCTCGCTCGAGTCTCTGAACGCTAGAGAGGTAAACATCGACCATACGTTGAGATC 364
QY 129 -----SerGlnProSerGln 133
Db 365 GTAGCTGGGAAAGTAGTGACGAGTAGAGTCTCTTGAATCGTTAGAGAAAGCTAGA 424
QY 134 GluIleValProGluThrSerAsnLysProGluGlyGly----- 146
Db 425 GATGGTTGTGTGATGTTCTTAATGATCATTCGCGTGTGATGTGACGATGAAGAGTCTT 484
QY 147 -----ArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIle 161
Db 485 CTCTCTGCTTTGAAATCTCTTTGTCTTAAATGATCGAGAGGTTTTTGGAACTTTGTG 544
QY 162 TyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLysLeu 181
Db 545 ACGCGGAGGAAGAGAGTGGGAAATCTCCGGTCAAGATTCGCGCGGCTTGGTTGAT 604
QY 182 AlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg 201
Db 605 TGTGTAGATCCGGCGATGTTAGTGTCTGAACGATATCTGAGGTTTTTCCGGTGGATACA 664
QY 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221
Db 665 AGAGGAGATAAAGTTAGCAATGATTATGATGGGCTTGTGTGTGATTCTTGAGAGTTTA 724
QY 222 GluSerPheLeuMetProAspArgGlyLysLysGlyLysIleGluSerTyrIle 241
Db 725 ACACCTGTAAATGTTGATCCAGTATTCGGAATTCGAGGTTGCTTACTCCGAGGTT 784
QY 242 LysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGly 261
Db 785 AAGGAGAAAGCTAAGAGATGCTCTGACACGTGGGAAGAAGATTTTGGGAAGAGAGGAGG 844
```

```
QY 262 LeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGly 281
Db 845 ATAGAGAATGTGACAGACTCTCTGATGTTTCATAGCTTTTTCGACGATCTTGTACATTTGGA 904
QY 282 ValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsn 301
Db 905 ATT-----GTTAAAGTGAAGATCTTGTGTTGTATAGAAAGCTT----- 943
QY 302 GluIleAlaGlyAlaLeuLysArgSerGlnPhe-----Leu 313
Db 944 ---GTTGTTGATCTCGCTGCGTAAACAGATCCCTAAAGTTGCGGTTTCTGTTGTTG 1000
QY 314 ValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAla 333
Db 1001 GGTGATCAATGCTGACATGATTGAGAAATTAATCAGCAGGGGCAACAGCTTGTATGCT 1060
QY 334 LeuGluMetValThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThr 353
Db 1061 GTTCATTTTACTTATGAAGTTGGGCTTGTGATAGTTCCTCCTGTTCTTGTCTCAA 1120
QY 354 SerPheLeuLysMetSerLysGluSerPheGluArg---AlaLysArgLysAlaGlnSer 372
Db 1121 GCTTATCTAAGGAGCGCAAGAAAGTCAGAGCTTCTATCATGGAGACTCTAGCAATACT 1180
QY 373 ProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGln 392
Db 1181 GCGCGAGCTACGATCTCGTGGCGGAGGCAATCAGCACCTTAAGCGGTTTTTAA 1240
QY 393 CysMetGluThrHisLysLeuAspProAlaLysGluLeuProGlyTyrGlnIleLysGlu 412
Db 1241 TGCATAGAAGAGTACAAACTCGAG-----GAAGAGTTCCCTCCAGAGAACTCAAGAAA 1294
QY 413 GlnIleValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAla 432
Db 1295 CGGTG-----GATCAGCTAGAGAA-GACCAAAACCGGAGAGAGAAAGCC 1338
QY 433 ArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnIle 452
Db 1339 AGCAGCTGTCTCTGCAACAAAGAACCG-----AGCCAGCTA 1377
QY 453 LysArg-ProArgLeuSerProMetGluMetProValThrSerSerSerTyrSerPr 472
Db 1378 CAACGTTCA-----ATGCCACCAAGCCGGCGGTATCACA 1419
QY 472 IleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAl 492
Db 1420 CGCATAGCTCTCT--TCCTTCCGTTTCATCAGATCACCTCTCACTCTCTCAATACGC 1476
QY 492 AlaValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSerPr 512
Db 1477 TTCACGAGCAGCTTACCCATCCCACTACACAGTC-----TACAGCAACAGAAAGCC 1530
QY 512 ogluTyr-----MetValProLeuProHisGlyLysLeuGlyArgSerVa 527
Db 1531 ACCTTATCCATACTCTCCGAGATCATCCAGGCTCATACCAAGCTCTCTATCGGTTA 1590
QY 527 lTyrAlaTyrGluHisLeu-----AlaProAsnSerTyrSe 539
Db 1591 CCCAGCGTACAAACGGTTATTGTAATGTCCTCAGTTCTCTGCTCCAGCTCTCCAGTTTACCA 1650
QY 539 rProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnAr 559
Db 1651 CCA---CACCATCATCA---CACCAACCATTC-----CACCATCAGA 1689
QY 559 gHis 560
Db 1690 ACAT 1693

RESULT 9
ADC03367
ID ADC03367 standard; DNA; 1377 BP.
XX
AC ADC03367;
```

XX 18-DEC-2003 (first entry)  
 XX DE Rice flowering time-related gene #4.  
 XX KW flowering time; rice; flower architecture; plant biotechnology;  
 KW commercial plant farming; agriculture; flowering-related protein; gene;  
 XX ds.  
 XX OS *Oryza sativa*.  
 XX PN WO2003000904-A2.  
 XX PD 03-JAN-2003.  
 XX PF 24-JUN-2002; 2002WO-BP006968.  
 XX PR 22-JUN-2001; 2001US-0300112P.  
 XX PR 26-SEP-2001; 2001US-0325277P.  
 XX PR 30-NOV-2001; 2001US-0334984P.  
 XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
 XX PI Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;  
 PI Glazebrook J, Katagiri F, Krens J, Provart N, Ricke D, Zhu T;  
 XX WPI; 2003-229340/22.  
 DR P-PSDB; ADC03368.  
 XX  
 CC New isolated nucleic acid molecule encoding a polypeptide modulating  
 PT flower architecture and flowering time, useful in the area of plant  
 PT biotechnology, and commercial plant farming and agriculture.  
 XX  
 PS Claim 4; SEQ ID NO 7; 323pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of proteins  
 CC involved in the control of flowering time in rice. The DNA and protein  
 CC sequences of the invention are useful for modulating flower architecture  
 CC and flowering time, the DNA and protein sequences are useful in the area  
 CC of plant biotechnology, commercial plant farming and agriculture. The  
 CC present DNA sequence encodes a rice flowering-related protein of the  
 CC invention.  
 XX  
 SQ Sequence 1377 BP; 392 A; 275 C; 352 G; 358 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3 47e-14 Length: 1377  
 Score: 293.00 Matches: 115  
 Percent Similarity: 42.11% Conservative: 85  
 Best Local Similarity: 24.21% Mismatches: 173  
 Query Match: 9.44% Indels: 102  
 DB: 9 Gaps: 19  
 US-09-890-475-1 (1-609) x ADC03367 (1-1377)  
 QY 36 ValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPhe 55  
 DB 10 ATGGAGTCCATGGCTACACTTATGAGTCCACA-----GGCTCCAAAGCTGCAGCAGCTT 63  
 QY 56 LeuLysSerIleAspGluLeuAlaAla----- 64  
 DB 64 CAGCGTGCAATTGCGAGGCTCAGAGCCAGCGCTGTTCTTGAACCTAAAGTGGAG 123  
 QY 65 -----PheSerValAlaValGluThrPheLysArgGlnPheAspAspLeu 79  
 DB 124 CAACTGGAGGACCAATTTCATGGGCTTGAGCAGCTCGCTCAAGAAAAGTTTGAATGATCTG 183  
 QY 80 GlnLysHisIleGluSerIleGluAsnAlaIle-----AspSerLysLeuGluSer 96  
 DB 184 AAAAGCAGAGAGAGGATTCAGAGGACCGGTTCGAAGTCAAGCATGCTGAGGCAA 243  
 QY 97 AsnGlyValValLeuAlaAlaArg----- 104

DB 244 CAAGAGGCTGTTGTCGTGGCAAGAACTGACTTCTTGTGAAAAAACTGCAGCAGAAAAGG 303  
 QY 105 -----AsnAsnAsnPheHisGlnProMetLeu 113  
 DB 304 GATGCGCGTGGCTGTGATCTTTGGCAAGTCCAAAGTGAATCTGCTACGCTCTTATC 363  
 QY 114 SerProProArgAsnAsnVal-----SerValGluThrThr 125  
 DB 364 AACCAATCAGCAAAATCCGTGAATAACAATGCTGTGTTTAAATGGAACATCGGTGAGT 423  
 QY 126 ValThrValSerGlnProSerGlnGluIleVal-----ProGlu 138  
 DB 424 CTTTCGTTAAGTGGCCAAACCTGCACTGGCATGGTGCATACCTGCAAGATGAAC 483  
 QY 139 ThrSerAsnLysProGluGlyArg-----MetCysGluLeuMetCysSerLysGly 156  
 DB 484 ACTGCTGTGAAGCCTCGTTCTCAACTTGTGCTCTGTGAAGAGATGAATGTAACGGG 543  
 QY 157 LeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIlePro 176  
 DB 544 CTTCAATAGTTCAATCAGACAAATCCAAAGACTTAAACATCCATCCGTGAGGAATCCCC 603  
 QY 177 SerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPhe 196  
 DB 604 GTTGCACTTAGGGGAGCAACTGATCATATGCTTGGTGTAGCTTCTCTAGAAGACTTC 663  
 QY 197 Tyr-----LeuGlnGlyArgAlaPheThrLysGluSerProMet 210  
 DB 664 TATTTGGAGACAACTTATATTGGATGGCAAA-----AAGATGGAACCTT 711  
 QY 211 SerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu---LeuMetProAsp 229  
 DB 712 CTGGGTGACGGAGGAGCGCTCCTGATGTAATCACTTCTCAGCTGCAGTGCAAACTGAT 771  
 QY 230 ArgGlyLysGlyLysValLys-----IleGluSerTrpIleLysAspGluAla 245  
 DB 772 GCTACTACTGTTTCATATCGGAGGTCAAGTCTTACTGCAAGATTAAGGCGGTGCA 831  
 QY 246 GluThrAlaAlaValAlaTrpArgLysArgLeu---MetThrGluGlyGlyLeuAlaAla 264  
 DB 832 AAAAGAGTGTCTTGAGTGGAAATCCAAAGTTGGATGAGCTTGACCTTGATGCTAGCAAT 891  
 QY 265 AlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer 284  
 DB 892 GGGAACTGCCTGGAAGCACATGCTTCTTCAACTCTCGCTACCTTGTATATTGCT 951  
 QY 285 AsnPheArgSerThrAspLeuLeuAspLeuIle-----ArgMetSerGlySer 300  
 DB 952 GAATTTGCACAGATGAGCTGTGTAAGTACTTCCCTCAGTCAGTCTGCTCGTCAGACA 1011  
 QY 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIle 320  
 DB 1012 CCTGAGCTTTGTAGAATACTTGGATTGTCCAGAACATG---CCAGGTGCTCATTTGGAGT 1068  
 QY 321 ValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuMetValTyrThrPhe 340  
 DB 1069 CTG-----GTTGAATGGAGAACATATTGATGCAATTAATTGGCTTATGCAATT 1119  
 QY 341 GlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys 360  
 DB 1120 GAGCTGACTAATCAGTTTGAACCGGTGAATTAATAAGACCTATCTACAGGAGGTCAAG 1179  
 QY 361 Glu-----SerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGlu 378  
 DB 1180 AGTGTGCCCATTTCAAGACTGGAAGATATCTCTGCAA-----GTACAGAAATGAA 1230  
 QY 379 AlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLys 398  
 DB 1231 ATGAATGAACGAGAGCTGTCTGCTTGAAGCTGCTATTAAAGTGTATCGAGGAGCACAAA 1290  
 QY 399 LeuAspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSerLeuGlu 418  
 DB 1291 CTTGATGAG-----AAATATCCGATTGACCTGCTTCAG 1323



PR	16-AUG-1999;	99US-0149368P.	PR	68	GAGATGGAGTCTTCCCGATCCAGGAGAGCTTGATTAAAGTCGCTCAGCCGAGTTCTTT	127																		
PR	17-AUG-1999;	99US-0149175P.	PR																					
PR	18-AUG-1999;	99US-0149426P.	PR																					
PR	20-AUG-1999;	99US-0149722P.	QY	57	LysSerIleAspGluLeuAlaIaIaPheSerValaIaValGluThrPheLysArgGln---	75																		
PR	20-AUG-1999;	99US-0149723P.	PR																					
PR	20-AUG-1999;	99US-0149929P.	PR	128	GAG-----	145																		
PR	23-AUG-1999;	99US-0149902P.	PR																					
PR	23-AUG-1999;	99US-0149930P.	PR																					
PR	25-AUG-1999;	99US-0150566P.	PR																					
PR	26-AUG-1999;	99US-0150884P.	PR																					
PR	27-AUG-1999;	99US-0151065P.	PR																					
PR	27-AUG-1999;	99US-0151066P.	PR																					
PR	27-AUG-1999;	99US-0151080P.	PR																					
PR	30-AUG-1999;	99US-0151303P.	PR																					
PR	31-AUG-1999;	99US-0151438P.	PR																					
PR	01-SEP-1999;	99US-0151930P.	PR																					
PR	07-SEP-1999;	99US-0152363P.	PR																					
PR	10-SEP-1999;	99US-0153070P.	PR																					
PR	13-SEP-1999;	99US-0153758P.	PR																					
PR	15-SEP-1999;	99US-0154018P.	PR																					
PR	16-SEP-1999;	99US-0154039P.	PR																					
PR	20-SEP-1999;	99US-0154779P.	PR																					
PR	22-SEP-1999;	99US-0155139P.	PR																					
PR	23-SEP-1999;	99US-0155486P.	PR																					
PR	24-SEP-1999;	99US-0155659P.	PR																					
PR	28-SEP-1999;	99US-0156458P.	PR																					
PR	29-SEP-1999;	99US-0156596P.	PR																					
PR	04-OCT-1999;	99US-0157117P.	PR																					
PR	05-OCT-1999;	99US-0157553P.	PR																					
PR	06-OCT-1999;	99US-0157865P.	PR																					
PR	07-OCT-1999;	99US-0158029P.	PR																					
PR	08-OCT-1999;	99US-0158232P.	PR																					
PR	12-OCT-1999;	99US-0158369P.	PR																					
PR	13-OCT-1999;	99US-0159293P.	PR																					
PR	13-OCT-1999;	99US-0159294P.	PR																					
PR	13-OCT-1999;	99US-0159295P.	PR																					
PR	14-OCT-1999;	99US-0159329P.	PR																					
PR	14-OCT-1999;	99US-0159330P.	PR																					
PR	14-OCT-1999;	99US-0159331P.	PR																					
PR	14-OCT-1999;	99US-0159637P.	PR																					
PR	18-OCT-1999;	99US-0159584P.	PR																					
PR	21-OCT-1999;	99US-0160741P.	PR																					
PR	21-OCT-1999;	99US-0160767P.	PR																					
PR	21-OCT-1999;	99US-0160768P.	PR																					
PR	21-OCT-1999;	99US-0160770P.	PR																					
PR	21-OCT-1999;	99US-0160814P.	PR																					
PR	21-OCT-1999;	99US-0160815P.	PR																					
PR	22-OCT-1999;	99US-0160980P.	PR																					
PR	22-OCT-1999;	99US-0160981P.	PR																					
PR	22-OCT-1999;	99US-0160989P.	PR																					
PR	25-OCT-1999;	99US-0161404P.	PR																					
PR	25-OCT-1999;	99US-0161405P.	PR																					
PR	25-OCT-1999;	99US-0161406P.	PR																					
PR	26-OCT-1999;	99US-0161359P.	PR																					
PR	26-OCT-1999;	99US-0161360P.	PR																					
PR	26-OCT-1999;	99US-0161361P.	PR																					
PR	28-OCT-1999;	99US-0161920P.	PR																					
PR	28-OCT-1999;	99US-0161992P.	PR																					
PR	28-OCT-1999;	99US-0161993P.	PR																					
PR	29-OCT-1999;	99US-0162142P.	PR																					
<p>Alignment Scores:</p> <table><tr><td>Pred. No.:</td><td>1,228-13</td><td>1865</td></tr><tr><td>Score:</td><td>288.50</td><td>132</td></tr><tr><td>Percent Similarity:</td><td>36.95%</td><td>103</td></tr><tr><td>Best Local Similarity:</td><td>20.75%</td><td>250</td></tr><tr><td>Query Match:</td><td>9.29%</td><td>151</td></tr><tr><td>DB:</td><td>3</td><td>21</td></tr></table>							Pred. No.:	1,228-13	1865	Score:	288.50	132	Percent Similarity:	36.95%	103	Best Local Similarity:	20.75%	250	Query Match:	9.29%	151	DB:	3	21
Pred. No.:	1,228-13	1865																						
Score:	288.50	132																						
Percent Similarity:	36.95%	103																						
Best Local Similarity:	20.75%	250																						
Query Match:	9.29%	151																						
DB:	3	21																						
<p>US-09-890-475-1 (1-609) x AAC38630 (1-1865)</p>																								
QY	37	GluThrGluSerThrSerMetAspIleThrLeuGlnSerLysGlnProGlnPheLeu	56																					

1079	AAAGCTTATCTGAGGGATGCAAGAAAGCTACAGCTTTTAATCACTGATGATGATTCACAAT	1138
366	AlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaIaThrLysGlnLeuAla	385
1139	TCTGGCGATCTGGCATCTT-----GTGGCGCGCAAGAGCAATCA	1180
386	ValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLysGluLeu	405
1181	GCACCTTAGACAGCTCTTTAAATGCATAGAAGAATACAACTCGAG-----GAAGATTC	1234
406	ProGlyTyrGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeuAsp	425
1235	CCACCCGAGAAATCTCAGAGAGAGGTTGGATCAGCTAGAGAAGACCAAAACCGAGAGAGA	1294
426	LysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaLeuAlaLys	445
1295	AAACGAGCGTGATTCGCCCAAC-----	1318
446	ArgMetTyrAsnGlnGlnIleLysArgProArgLeuSer---ProMetGluMetProPro	464
1319	-----AAGAGAAACCGAGCAAGCTACAGCGGTCCAATGCCACCA	1357
465	ValThrSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArgAsp	484
1358	GCCAAAGACGGGGGTATCACA-----	1378
485	AspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPhe	504
1379	-----AAGCGTTACGTTCTCCTCTTTCCACCGCGCTCCACCAACATTC	1420
505	ProHisArgSerArgArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGly	524
1421	ATCAGATCCCAAAGTCACTCACTCAATACGGGTATCCA-----	1459
525	ArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyr-----SerProGly	541
1460	-----GCATACACCATATCCCCACCTACCATCTACACCAACAGAGGCCCTCCA	1507
542	HisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGly-----GlnArg	559
1508	TAC-----CAATACTACCCGAGCGGTTTCATGGTTCTTACCAAAACC	1549
560	HisProLeuGlnTyrSerProIleHisGlyGlnGlnLeuLeuProTyrGlyIle---	578
1550	TCCTCCGTCAGTTAT---CCGACAGCATATGGTACTACTGCACTCCGGTGGGTGCTCCA	1606
579	-----GlnArgValTyrArgHisSerProSerGluGluArgTyrLeu	592
1607	CCCCCTCCAGTTTACCATCTCCACCCGACCAACCAACCATATTCAGCATGCTTACTAC	1666
593	GlyLeuSerAsnGlnArgSerProArgSerAsnSerSerLeuAspPro	608
1667	TGATTACCAATGACAAAGGAAAAAAGATGATGAAGTACCAACCA	1714

	24-AUG-2000; 2000US-0227866P.	
	26-JAN-2001; 2001US-024647P.	
	22-JUN-2001; 2001US-030011P.	
XX	(SCRI ) SCRIPPS RES INST.	
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.	
XX		
PI	Harper JF, Kreps J, Wang X, Zhu T;	
XX		
DR	WPI; 2002-304127/34.	
XX		
PT	Identifying a stress condition to which a plant cell has been exposed and	
PT	producing plants with increased tolerance to these abiotic stresses.	
XX		
PS	Claim 144; SEQ ID NO 917; 577pp + Sequence Listing; English.	
XX		
CC	The invention relates to identifying a stress condition to which a plant	
CC	cell has been exposed, comprising: (a) contacting nucleic acid	
CC	representative of expressed polynucleotides in the plant cell with an	
CC	array or probes representative of the plant cell genome; and (b)	
CC	detecting a profile of expressed polynucleotides in the plant cell	
CC	characteristic of a stress response. The method is useful in the	
CC	production of transgenic plants, cells and seeds and in producing plants	
CC	with increased tolerance to abiotic stress. The present sequence is that	
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used	
CC	in methods of the invention. Note: The sequence data for this patent is	
CC	not represented in the printed specification but is based on sequence	
CC	information supplied to Derwent by the European Patent Office	
XX		
SQ	Sequence 1524 BP; 401 A; 342 C; 391 G; 390 T; 0 U; 0 Other;	
 Alignment Scores:		
Pred No.:	2,25e-09	Length: 1524
Score:	234.50	Matches: 117
Percent Similarity:	34.38%	Conservative: 91
Best Local Similarity:	19.34%	Mismatches: 222
Query Match:	7.55%	Indels: 175
DB:	6	Gaps: 21
 US-09-850-475-1 (1-609) x ABZ13112 (1-1524)		
Qy	49 GluSerIysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPheSerValala 68	
Db	34 AAGTCGTCTCAGCGGATTTCTTTTGAG-----PheAspAsp 78	
Qy	69 ValGluThrPheLysArgGln-----PheAspAsp 78	
Db	61 -----TTTCAAGACGGGCTCTTAATCACGAGCTGCAATCTTCTATGGAAAGAG 111	
Qy	79 LeuGlnYHisIleGluSerIleGluAsnAlaIleAspSerIlysLeuGluSerAsnGly 98	
Db	112 CTCCTCAGCATTTCACTTCGCATGGAGCAAACCTGATGAAGAATACTGAAGCTCTGAGA 171	
Qy	99 ValValLeuAlaAlaArgAsnAsnAsnPheHISglnProMetLeuSerProProArgAsn 118	
Db	172 CAGATGATTGAACCCCTAGATATCAAACCCAGACGCTCGATTGGTTACTCAAGCATCGT 231	
Qy	119 AsnValSerValGluThrThrValThrValSer----- 129	
Db	232 GAGGTGACTATAGACCACGCGTCGAGATCCGGAGGGGAAAGTTGAGGAACGCGTTTAGA 291	
Qy	130 -----GlnProSerGlnGluIleValProGluThrSerAsnLysPro 143	
Db	292 GCGGCGTTGATTGCTGGAGAAAGCTAGAGACTGTGGTGATGAAGATACCGGAGAAGTC 351	
Qy	144 GluGlyGly-----ArgMetCysGluLeuMetCysSerLys 155	
Db	352 GACGACGGCAGCGGCTTTTATCGGCCCTGAAGTCGTTTTGTTGAAGATGGACCAAGA 411	
Qy	156 GlyLeuArgLysTyrlleTyAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIle 175	
Db	412 CGATTTTGGGGGTTTGATTTGCCGAGAAGAAGATTTGGGAATCTCCGGTCCACAGTT 471	

Qy 515 MetValProLeuProHisGlyGlyLeuGlyArgSerValTyAlaTyTyrGluHisLeuAla 534  
| | | | |  
Db 1306 GGTGTACCA-----GCATACACCACATCCCCA 1332  
| | | | |  
Qy 535 ProAsnSerTyr-----SerProGlyHisGlyHisArgLeuHisArgGlnTyrSer 551  
| | | | | : : : : :  
Db 1333 CCTACCATCTACAGCAACAGGAGCCCTCCATAC-----CAATACTCA 1374  
| | | | |  
Qy 552 ProSerLeuValHisGly-----GlnArgHisProLeuGlnTyrSerProProlleHis 569  
| | | | | : : : : :  
Db 1375 CCGAGGGCGGTTTCATGGTCTTACCAAACCTCTCCGTCAGTTAT---CCGACAGCATAT 1431  
| | | | | : : : : :  
Qy 570 GlyGlnGlnGlnLeuProTyrGlyIle-----GlnArgValTyrArgHisSerProSer 587  
| | | | | : : : : :  
Db 1432 GGTACTACTGCAGTCCGGTGCTGCCACCCCCTCCAGTTTACCATCTCACCOCGCAC 1491  
| | | | | : : : : :  
Qy 588 GluGluArgTyrLeu 592  
| | | | | : : : : :  
Db 1492 CACACACCATTAT 1506  
| | | | | : : : : :  
  
RESULT 12  
ID ADC03447  
AD ADC03447 standard; DNA; 947 BP.  
XX AC ADC03447;  
XX DT 18-DEC-2003 (first entry)  
XX DE Wheat flowering time-related DNA sequence #7.  
XX KW flowering time; wheat; flower architecture; plant biotechnology;  
KW commercial plant farming; agriculture; flowering-related protein; ds.  
XX OS Triticum aestivum.  
XX FN WO2003000904-A2.  
XX PD 03-JAN-2003.  
XX PF 24-JUN-2002; 2002WO-EP006968.  
XX PR 22-JUN-2001; 2001US-0300112P.  
PR 26-SEP-2001; 2001US-0325277P.  
PR 30-NOV-2001; 2001US-0334984P.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX PI Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;  
PI Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;  
XX DR WPI; 2003-229340/22.  
XX PT New isolated nucleic acid molecule encoding a polypeptide modulating  
PT flower architecture and flowering time, useful in the area of plant  
PT biotechnology, and commercial plant farming and agriculture.  
XX PS Disclosure; SEQ ID NO 87; 323pp; English.  
XX CC The invention comprises the amino acid and coding sequences of proteins  
CC involved in the control of flowering time in rice. The DNA and protein  
CC sequences of the invention are useful for modulating flower architecture  
CC and flowering time, the DNA and protein sequences are useful in the area  
CC of plant biotechnology, commercial plant farming and agriculture. The  
CC present wheat DNA sequence shows homology to the rice sequences of the  
CC invention.  
XX SQ Sequence 947 BP; 248 A; 229 C; 231 G; 239 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 947  
Score: 224.00 Matches: 82  
Percent Similarity: 42.14% Conservative: 52  
Best Local Similarity: 25.79% Mismatches: 100

QY	176	ProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValIleAspCysIleGlyLys	195
DB	472	CCTGTGGCGTGGTAGATTGTGTGGATCCGCCGAAGTTAGTACTTGAAGCTGATCTGAG	531
QY	196	PheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAlaargLn	215
DB	532	GTTTTTCGGTGGAATAAAGAGGT--GGTGGAGAGAAGTGGAGCAATGATTTTGGATGG	588
QY	216	ValSerLeuLeuIleLeuGluSerPhe-----LeuLeuMetProAspargLyLys	232
DB	589	GCTTGTGTGTGATTCTGGAGAGTTTGATTCACGTTTATGTTGATCCAGTGATGGGAAG	648
QY	233	GlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaLaValalaTrp	252
DB	649	TCTCGGCTACTTGTAATCCTACGGTTAAAGGAAAAGCTAAAGAGATTGCTGAGACGTGG	708
QY	253	ArgLysArgLeuMetThrGluGlyGlyLeuAlaLaalagLuLysMetAspAlaargLy	272
DB	709	AAGCTAGCTTGGAAAGAGAGAGGAGGATGTGAGAATGTGAAAAACACTGATGTTCATACG	768
QY	273	LeuLeuLeuLeuValLalaCysPheGlyValProSerAsnPheArgSerThrAspleuLeu	292
DB	769	TTTCTGCAACATCTGTGCATTTTTGGGATTGTGAAGAAG-----GATGATCTA	816
QY	293	AspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPhe	312
DB	817	GCTCTTTATAGG-----AAACTGTGTGTTGGTTCAGCTTGGCGTAAACAGATG	864
QY	313	-----LeuValProMetValSerGlyIleValGluSerSer	324
DB	865	CCAAAGCTTGCTGTTTCAGTTGGTTGGCGCACCAAATGCCTGATGATTGAAGAGTTG	924
QY	325	IleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp	344
DB	925	ATAATCAGGGGACACACAGCTTGATCGGTTCAITTCATTTCAAGTTGGTCTGTATCAC	984
QY	345	LysPheSerAlaLaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGlu	364
DB	985	TTGTTCCTCTCTGTTCTTTCCTCAAAGCTATTCTGAGGGATCCAAAGAAGACTACAGCT	1044
QY	365	-----ArgAlaLysArgLysAlaGlnSerProLeu	374
DB	1045	TTAATCACTGATGATCCAAATTTGCGCGGATCTCGGGATCTGTTCTTCGCACTC	1104
QY	375	AlaPheLysGluAlaLaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMet	394
DB	1105	GCATTAC-----CAATGTATG	1119
QY	395	GluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIle	414
DB	1120	TGCTAGAGAGACCAAAACCGAGAGAGAAACCAGCC-----GTG	1161
QY	415	ValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSer	434
DB	1162	ATTCGCCGCAACAAGAGAAC-----	1182
QY	435	LeuSerLeuMetGluGluAlaLaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArg	454
DB	1183	-----CGAGCAAGCTACAGCGT-----	1200
QY	455	ProArgLeuSerProMetGluMetProProValThrSerSerSerTyrSerProIleTyr	474
DB	1201	-----CCAATGCCACCGACCAAGCAGCGGCGTATCAC-----	1233
QY	475	ArgAspArgSerPheProSerGlnArgAspAspGlnaspGluIleSerAlaLeuVal	494
DB	1234	-----AACGTTACGTC	1245
QY	495	SerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgArgSerProGluTyr	514
DB	1246	TCCCTTTCCCACCGCGCTCCACCAACATTCATCAGATCCCCAAAGCTCACTCACCTCAATAC	1305

RESULT 13  
ABL17653  
ID ABL1  
XX

Db 3709 TTCAGGGAAG-----CCGGAACCTGGAGTACGAACCGAT 3744  
QY 92 SerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsnHeHisGlnPro 111  
Db 3745 CCGGAGTAATGCGT-----ATCTCAAGAAGAGCCCATAGCACAAATGAACCA 3795  
QY 112 MetLeuSerProProArgAsnAsnValSerValGluThrThrValThr-----ValSerGln 130  
Db 3796 -----TCTGACCGCGCAGAGATAAGTCAACCGATGATGTAAGAACTTTGTATGGCAA 3849  
QY 131 ProSerGlnGluLeuValPro-----GluThrSerAsnLys 142  
Db 3850 CACATTCAGAAATTTGCCCAACCTGTCGCAACAGGTTGAACAGAGCTCATCGAGAA 3909  
QY 143 ProGluGlyArgMetCysGlnLeuMetCysSerLysGlyLeuArgLysTyrIleTyr 162  
Db 3910 CCGGAAGAGGTAGCTAAATTTAG----- 3933  
QY 163 AlaAsnIleSerAspGlnAlaLys-----LeuMetGluGluIleProSerAla 178  
Db 3934 -----CGAAGACAGCCAAAGAGGCTGAACCTCTACTCGAACCTCAGCCCTGAGCCC 3984  
QY 179 LeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeu 198  
Db 3985 ACACCA--AGTGTGATCTCTGCTCTACATCATGGAGGAAGTCATCAAGCCCAATTGA 4041  
QY 199 GlnGlyArgArgAlaPheThrLysGluSerProMetSerSer-----AlaArgGln 215  
Db 4042 -----AGGGAAGCCAGCATTCGTGAGGAAGTTCAACCCAGTTTATCTATGCACTGAA 4095  
QY 216 ValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysVal 235  
Db 4096 ATAGCC-----AATCTCAAGTTGGAATTTGATCGTGGC----- 4128  
QY 236 LysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTTPArgLysArg 255  
Db 4129 ---ACTGAACGATGG-----CACAGGCGGAA-----TGGGAAGAGCAGC 4164  
QY 256 LeuMetThrGluGlyGlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeu 275  
Db 4165 GATCTTACGGAT-----GCAGAACGC-----ATA 4188  
QY 276 LeuValAlaCysPheGlyValProSerAsnAsnAsnSerThrAspLeuLeuAspLeuIle 295  
Db 4189 GTATCAAAATGTTATATGCTCTTAATGATGCAACCAAGGAGGATGTTGTAGATGTACGC 4248  
QY 296 ArgMetSerGlySerAsnGluIle-----AlaGlyAlaLeuLysArgSerGlnPheLeu 313  
Db 4249 GTTGAAGTTCACTCAATTCGCTTCCACAGCAAGAAAGCACAGAGAATCTGTCTCCAGAA 4308  
QY 314 ValProMetValSerGlyValGluSerSerIleLysArgGlyMetHisIleGluAla 333  
Db 4309 GTACCAATTGAAGCTCAAAAGATTGAAATACA-----GAAAGTTACACAA 4353  
QY 334 LeuGluMetVal---TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeu 352  
Db 4354 CCAGAGACCGTTAATGAGAACTTCGCGAGCAAGAAATCTAGAACACCATTTAGTTAT 4413  
QY 353 ThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSer 372  
Db 4414 ACAAAAATCATTAAATATAGCGCTAGTACTCATCAATATCCAAAGAAACAAAGCAAAAGT 4473  
QY 373 ProLeuAlaPheLysGluAlaAlaThrLysGlnIleLeuAlaValLeuSerSerValMetGln 392  
Db 4474 CCGAAGCGCTTTTCAAGATTCCCGTAAGAACCCCTGAGCAGCAGCAGTCTTCGCGAGTGC 4533  
QY 393 CysMetGluThrHisLysLeuAspProAlaLysGlu----- 404  
Db 4534 ACTCGTTCGCAATCGACTCCACGCGCCATGATGATCGAAGGGAAGAGACCAACG 4593  
QY 405 LeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeu 424  
Db 4594 TCACAAGGTGTACCATAGGGGAAACAGTCAGTTCTCCAAATCCGAACAGTGTGTCGAC 4653

QY 425 AspLysGluMetGlu-----GluLysAlaArgSerLeuSerLeuMet 438  
Db 4654 AACCCAGAGGTTAATCTTGTGAGCCCAAGAAACACACATGCAAAAAGGATACAAATATTGTC 4713  
QY 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSer 458  
Db 4714 GAAGAACCTGCCACCCAGCCTTTAGCTTAGAGCTGGAGAACATAGTCCGATCAACT 4773  
QY 459 ProMetGluMetProProVal-ThrSerSerSerTyrSerProIleTyrArgAspArgSe 478  
Db 4774 GCTGTAGCGTTCAGTCCACCGACTCGGACGAAGATTTCGAAGATCTCTGAGTTCAGC 4833  
QY 478 rPheProSerGlnArgAspAsp-----AspGlnAspGluIleSerAlaLeu 493  
Db 4834 GTTAGCATGGAAACAGACCCCATGATGAACCCGCTTCGATGACAGCTTTATAGCTT 4893  
QY 493 uValSer-----SerTyrLeuGlyProSerThrSerPheProHisArgSerArg----- 509  
Db 4894 GATAGCATGGCAGCGGCTCAAAACCAAGT-----CCACAGAGGATGAAGTCGTA 4944  
QY 510 -----ArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGlyArg 525  
Db 4945 ATCTTCTCGATGAGATCCAGATGGAATCTTCGATAGCCCC----- 4987  
QY 525 gSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArg 545  
Db 4988 -AGTCCGAAGCAACGCAAGTCTTGATCCCACTCAAGTGAATCCG-----AG 5034  
QY 545 gLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSe 565  
Db 5035 ACCTCCAAAGTTGTTCTCAAGGAGTTG-----TACCCTCTCGAGATCCG 5079  
QY 565 rPro 566  
Db 5080 GCCA 5083  
RESULT 14  
AAZ32435  
ID AAZ32435 standard; cDNA; 3312 BP.  
XX AAZ32435;  
AC AAZ32435;  
XX 27-JAN-2000 (first entry)  
DE DE Rat TAO1 kinase encoding cDNA.  
XX TAO1; TAO2; MEK3; mitogen activated protein kinase; phosphorylation; p38;  
XX protein kinase; cancer; inflammation; autoimmune disease; degeneration;  
XX insulin-resistant diabetes; metabolic disorder; neurodegeneration;  
XX MAP kinase; MAP/ERK kinase; ss.  
OS Rattus sp.  
XX Key Location/Qualifiers  
XX 121..3126  
XX /tag= a  
XX /product= "TAO1 protein kinase"  
XX WO9953076-A1.  
XX 21-OCT-1999.  
XX 14-APR-1999; 99WO-US008165.  
XX 14-APR-1998; 98US-00060410.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX Cobb M, Hutchison M, Chen Z, Berman K;  
XX WPI; 1999-633831/54.  
XX P-PSDB; AAY49896.



XX New polypeptides that phosphorylate kinase, used to screen for modulators  
PT for treating e.g. cancer or inflammation.

XX Claim 12; Fig 1; 95pp; English.

CC The present sequence encodes rat TAO1 protein kinase, which is capable of  
CC phosphorylating MEK3 (a MAP/ERK kinase). TAO kinases, and related  
CC polypeptides, are used to screen for modulators of stress-responsive  
CC mitogen activated protein (MAP) kinase pathways. These modulators are  
CC potentially useful for treating or preventing: (1) inflammation,  
CC autoimmune disease, cancer and degeneration (inhibitors of  
CC phosphorylation); or (2) insulin-resistant diabetes, metabolic disorders  
CC and neurodegeneration (enhancers of phosphorylation). TAO kinases are  
CC also used to raise specific antibodies, useful therapeutically as  
CC modulators and as immunoassay reagents for detecting TAO kinases. TAO  
CC kinase polynucleotides can be used: (a) for recombinant expression of TAO  
CC kinases; and (b) in the form of fragments, for detecting TAO kinase  
CC polynucleotides in standard hybridisation and amplification tests. TAO  
CC kinases are highly specific for MEK3

XX Sequence 3312 BP; 1106 A; 674 C; 780 G; 752 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 0.0391 Length: 3312  
Score: 151.00 Matches: 150  
Percent Similarity: 34.75% Conservative: 95  
Best Local Similarity: 21.28% Mismatches: 237  
Query Match: 4.86% Indels: 223  
DB: 2 Gaps: 32

US-09-890-475-1 (1-609) x AAZ32435 (1-3312)

QY 3 AsnTyr-----ProProThrValAlaAlaGlnProThrThrThrAlaAsn 17  
DB 1333 AATTACCAGNAGGAGGAGATCTTAGAACAGAGCATCAGCTCCACAGTCTCCA-----1386  
QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGlnLeuProLysIleValGlu 37  
DB 1387 CCTCAAGTGTCTGCTCAAAATCACAFTATCGTAATAGAACACACTTGTGCAACTATACGA 1446  
QY 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57  
DB 1447 ACAGCATCACTG-----GTTACAGACAGATGCAAGAACATGAG 1485  
QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77  
DB 1486 CAGGACTCTGAACCTTAGAGAACAGATCTCTGTTATAAGCGGATGAGCGACAGCAT---1542  
QY 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97  
DB 1543 -----CAGAACGAGCTGATGACTCTGTGAAATTAACCTGAAGGAGGAGAAATGGACGACAT 1596  
QY 98 GlyVal-----ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113  
DB 1597 CGGCTCAGATTAGACAAAGATCTTGAACCTCAGCGCAACAATTTGCTGCGAGAAATG---1653  
QY 114 SerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133  
DB 1654 -----GAGAACTTATTAGAAACACCAAGCTTCTATG 1686  
QY 134 GluIleValProGluThrSerAsnLysProGluGlyArgMetCysGluLeuMetCys 153  
DB 1687 GAAAGAGGCTAAAGTGATGGCCCAACGAGGAG-----1719  
QY 154 SerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173  
DB 1720 -----AAAAAATTCACACACAT-----CAGGCTCAACAGAGAAA 1758  
QY 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysIle 193  
DB 1759 GAACCTGAATAGCTTTTGGAGTCTCAAAAGAGAGATATAAAGTCTGAAAAGAGACAGCTT 1818

QY 193 eGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSer-----ProMetse 211  
DB 1819 AAGGAGGAGCTGAATGAAGAACACAGGACACACCTTAAAGAAAGAGAGGAAATGGCTTTCA 1878  
QY 211 rSerAlaArgGlnValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 231  
DB 1879 AAGCAGAGAGGAGATATTCAACATTTTCAGGCAAGAGAGAGATATCTTCTTCGACGT 1938  
QY 231 yLysGlyLysVal-----LysIleGluSerTyrIleLysAspGluAlaGluTh 247  
DB 1939 CAAGGAGGATCTAGAGCTAGAGTATGTCGCTTCA-----AAAGAGAAATGTTACTT 1992  
QY 247 rAlaAlaValAlaTyrArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267  
DB 1993 GGTGGCATACTTGGAAACAGGACCTTG-----TCAGGGAGGAGTT-----AAACAAA 2040  
QY 267 yMetAspAlaArgGlyLeuLeuLeu-----LeuValAlaLac 279  
DB 2041 AGGAGAGCTCAGAGGAGCTTAGAACATGCAATGTTACTGGACAGCATGAATCATGCAA 2100  
QY 279 yPheGlyVal---ProSerAsnPheArgSerThrAspLeuLeu-AspLeuIleArgMet 297  
DB 2101 GAACCTGGAGTTTCGCCACCTCAACACTATTTCAGAGAGATGGCTGTGATTCAGACTG 2160  
QY 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317  
DB 2161 CAACATCAAACTGAGCTTACTTAACAGCTGGAATACAAATAGAGAGAG-----2208  
QY 318 SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal 337  
DB 2209 -----GAACGGGAACTAAGCGG---AAACATGTC-----2235  
QY 338 TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357  
DB 2236 -----ATGGAAGTTCGACAGCAGCTAAGAGTTTGAAGTCTTAAAGAACTCCAA 2283  
QY 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla-----375  
DB 2284 ATAAAAAGCAGTTTCAGGATACCTGCAAAATTCAAACCAGACAGTCAAAAGCATTAAAG 2343  
QY 376 -----PheLys 377  
DB 2344 AATCACTACTGGAGACTACACCAAGAGTGAAGCAAAAGCTGTCTCTGAAAGACTCAAG 2403  
QY 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392  
DB 2404 GAGGAACAGACTCGGAAGTATGACCATCTTGGCTGAGCAGTATGATCATAGCATTAATGAA 2463  
QY 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu-----404  
DB 2464 ATGCTCTCCACACAGCTCTGCTTGGATGAAGCAGCAGGAGAGAGATGCCAGGTTTG 2523  
QY 405 -----LeuProGlyTyrTrpGlnIleLys-----411  
DB 2524 AAGATGCAGCTACAGCAGAACTGGAGCTGTGAATGTCATATCAGACAAATCAAGATG 2583  
QY 412 -----GluGlnIleValSerLeu-----417  
DB 2584 CAGGCTGAGGCCCAACATGATCGAGCTTCGAGAGCTGGAACAAAGGTTCTCCCTTCGG 2643  
QY 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429  
DB 2644 AGAGCACTCTTAGAACAGAGATTTGAAGAGAGATGTTGGCTTTTCAGAGATGAACGCCACA 2703  
QY 430 GluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsn 449  
DB 2704 GAACGAATACGT-----AGCTGCTGAGCGCCAGGCGCAGAGAAATGAAGCTTTTGAC 2757  
QY 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSer 469  
DB 2758 TCTGAAAGCATGAGATTAGGTTTATGAATACATGGTCTT-----TCTAAT 2802  
QY 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489

```
Db 2803 CTCTCCCTCAGGCATTGAGCCACAGCTACCCAGGA----- 2838
Qy 490 IleserAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509
Db 2839 -----CCTTCTAGCTGGTCTCACAAT----- 2859
Qy 510 ArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAla 529
Db 2860 -----CCTACTGGGGGTTACGGA----- 2877
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549
Db 2878 -----CCTCACTGGGGTCAATCCCATG----- 2898
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566
Db 2899 -----GTTGGCACACCAAGCTTGGGGTCAATCCGATGCAAGCGGACCCCAACCATGG 2952
Qy 567 -----ProIleHisGlyGlnGlnLeuProTyrGlyGlyGlnArgValTyrArgHis 584
Db 2953 GGTCACTCCCTCAGGGCCCAATCAAGGGGTACCTCGAGTAGCAGTATAGGATCCGCAAT 3012
Qy 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerAsnGlnArgSerProArg 601
Db 3013 AGCCCCAGGCTCTCAGGCGGACAGCTTCTGGGGGCGGACGGAACAGGGCATGAGCAGA 3072
Qy 602 SerAsnSer 604
Db 3073 AGCAGAGT 3081
```

## RESULT 15

ABT40852  
ID ABT40852 standard; DNA; 3312 BP.

XX AC ABT40852;

DT 26-JUN-2003 (first entry)

XX DE Toxicity modelling related rat gene SEQ ID No 554.

XX KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;  
KW database; drug screening; toxicity assay; rat; ds.

XX OS Rattus norvegicus.

XX PN WO200295000-A2.

XX PD 28-NOV-2002.

XX PF 22-MAY-2002; 2002WO-US016173.

XX PR 22-MAY-2001; 2001US-0292335P.

XX PR 13-JUN-2001; 2001US-0297523P.

XX PR 19-JUN-2001; 2001US-029825P.

XX PR 10-JUL-2001; 2001US-0303807P.

XX PR 10-JUL-2001; 2001US-0303808P.

XX PR 28-AUG-2001; 2001US-0315047P.

XX PR 27-SEP-2001; 2001US-0324928P.

XX PR 22-OCT-2001; 2001US-0330462P.

XX PR 01-NOV-2001; 2001US-0330867P.

XX PR 06-DEC-2001; 2001US-0336144P.

XX PR 19-DEC-2001; 2001US-0340873P.

XX PR 21-FEB-2002; 2002US-0357842P.

XX PR 21-FEB-2002; 2002US-0357843P.

XX PR 15-MAR-2002; 2002US-0357844P.

XX PR 08-APR-2002; 2002US-0364134P.

XX PR 08-APR-2002; 2002US-0370206P.

XX PR 08-APR-2002; 2002US-0370247P.

XX PR 17-APR-2002; 2002US-0372794P.

PR 21-APR-2002; 2002US-0371679P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

XX XX WPI; 2003-148464/14.

XX DR Predicting at least one toxic effect of a compound, useful for toxicity  
PT modelling, comprises preparing a gene expression profile of a tissue or  
PT cell sample exposed to the compound, and comparing the gene expression  
PT profile to a database.

XX PS Example 4; Page; 446pp; English.

XX CC The invention relates to a novel method of predicting at least one toxic  
CC effect of a compound. The method comprises a gene expression profile of a  
CC tissue or cell sample exposed to the compound, and comparing the gene  
CC expression profile to a database comprising at least part of the data or  
CC information given in the specification. The methods are useful for  
CC predicting at least one toxic effect of a compound, predicting the  
CC progression of a toxic effect of a compound, predicting the renal  
CC toxicity of a compound, or identifying toxicity markers in tissues or  
CC cells exposed to known renal toxin. The genes are useful as toxicity  
CC markers in drug screening and toxicity assays, in monitoring disease or  
CC physiological states, or disease progression. This polynucleotide  
CC represents a rat DNA sequence relating to the toxic effect database  
CC described in the specification. NOTE: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the World Intellectual Property  
CC Organization

XX SQ Sequence 3312 BP; 1106 A; 674 C; 780 G; 752 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0.0391 Length: 3312  
Score: 151.00 Matches: 150  
Percent Similarity: 34.75% Conservative: 95  
Best Local Similarity: 21.28% Mismatches: 237  
Query Match: 4.86% Indels: 223  
DB: 7 Gaps: 32

US-09-890-475-1 (1-609) x ABT40852 (1-3312)

Qy 3 AsnTyr-----ProProThrValAlaAlaGlnProThrThrThrAlaAsn 17

Db 1333 AATTACCAAGAAGAGGAGATCCTAGACACAGAGCATCAGTCCACAGTCTCCA----- 1386

Qy 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37

Db 1387 CCTCAAGTGTCTCGTCAAAATCAATATCTGTAATAGAGAACACTTTGCACTATACGA 1446

Qy 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57

Db 1447 ACAGCATCACTG-----GTTACAAGACAGATGCAAGACATGAG 1485

Qy 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77

Db 1486 CAGGACTCTGAATCTAGAACACAGATGCTGTGTTATAGCGGATGAGGCGCAGCAT--- 1542

Qy 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97

Db 1543 -----CAGAGCAGCTGATGATCTCGGAAATATAACTAAGCGAGAAATGCGCAACAT 1596

Qy 98 GlyVal-----ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113

Db 1597 CGGCTCAGATTAGACAAAGATCTTGAACACTCGCGCAACAAATTCGCTCGCAAAATG--- 1653

Qy 114 SerProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133

Db 1654 -----GAGAACTTATTAAAGAAACACCAAGCTTCTATG 1686

Qy 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153

Db 1687 GAAAAAGAGCTAAAGTGTATGGCCACAGCAGAG----- 1719  
Qy 154 SerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173  
Db 1720 ---AAAATTCACACACATTT-----CAGGCTCAACACAGAGAA 1758  
Qy 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysII 193  
Db 1759 GAACGTGAATAGCTTTTGGAGTCTCAAAAAGAGAAATATAAATCTCGAAAGAGCAGCTT 1818  
Qy 193 eGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSer-----PrometSe 211  
Db 1919 AAGCAGAGCTGAATGAACACAGCAGCACACCTTAAAAAGAAAGCAGGATGGCTTCA 1978  
Qy 211 rSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGlu 231  
Db 1879 AAGCAGAGGAGATATTCAACATTTTCAGGCACAGAGAAAGCTATCTTCTTCGACGT 1938  
Qy 231 YLysGlyLysVal-----LysIleGluSerTrpIleLysAspGluAlaGluTh 247  
Db 1939 CAAGGAGATCTAGAGCTAGATGTCTGCGTCTCA-----AAGAGAAATGTACTT 1992  
Qy 247 rAlaAlaValAlaTrpArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267  
Db 1993 GGTCCGATTAATTTGGAACAGGACCTTG-----TCAGGAGGAGTT-----AAACAAA 2040  
Qy 267 yMetAspAlaArgGlyLeuLeuLeu-----LeuValAlaLac 279  
Db 2041 AGGCAGACTCAGAGGACTTAGAACATGCATGTACTTGGCAGCAGCATGAATCCTGCAA 2100  
Qy 279 yPheGlyVal---ProSerAsnPheArgSerThrAspLeuLeu-AspLeuIleArgMet 297  
Db 2101 GAACGTGGAGTTTCGCCACCTCAACACTATTTCAGAGATGGCTGTGAGTTGATCAGACTG 2160  
Qy 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValPrometVal 317  
Db 2161 CAACATCAAACTGAGCTTACTTAACACGCTGGAATACATAAGAGAGG----- 2208  
Qy 318 SerGlyIleValGluSerSerIleLysArgGlyMethIleGluAlaLeuGluMetVal 337  
Db 2209 -----GAACGGAACTAAAGCGG---AAACATGTC----- 2235  
Qy 338 TyrThrPheGlyMetGluAspLysPheSerAlaLeuValLeuThrSerPheLeuLys 357  
Db 2236 -----ATGGAAGTTTCGACAGCAGCTTAAGAGTTTGAAGTCTAAAGAACTCAA 2283  
Qy 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375  
Db 2284 ATAAAAAGCAGTTTCAGGATACCTGCANAAATTCAAACCAGACAGTACAAAGCATTAAGG 2343  
Qy 376 -----PheLys 377  
Db 2344 AATCACCTACTCGAGACTACACCAAGAGTGAGCACAAGAGTGTCTGAAAAGACTCAAG 2403  
Qy 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392  
Db 2404 GAGGAACAGACTCGGAAGTTAGCCATCTTGGCTGAGCAGTATGATCATGATTAATGAA 2463  
Qy 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu----- 404  
Db 2464 ATGCTCTCCACACAAGCTCTCGGTTGGATGAAGCAGGAGCAGAGTGCAGGTTTGG 2523  
Qy 405 -----LeuProGlyTrpGlnIleLys----- 411  
Db 2524 AAGATGCAGCTACAGCAGGAATCGAGCTGTGTAATGCATATCAGACAAATCAAGATG 2583  
Qy 412 -----GluGlnIleValSerLeu--- 417  
Db 2584 CAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTGGACAAAGGGTCTCCCTTCG 2643  
Qy 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429

Db 2644 AGACACACTCTTAGACACAGAGATTCAAGAGAGATGTTGGCTTTGCAGATGAAGGCACA 2703  
Qy 430 GluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsn 449  
Db 2704 GAACGAATACCT-----AGCCTGCTCGAGCCGACAGAGAAATTTGAAGCTTTTGAC 2757  
Qy 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSer 469  
Db 2758 TCTGAAAGCATGAGATTAGGTTTAACTAGTACATGGTCTT-----TCTAAT 2802  
Qy 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489  
Db 2803 CTCTCCCTCAGGCATTCAGCCACAGCTACCCAGGA----- 2838  
Qy 490 IleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509  
Db 2839 -----GCTTCTAGCTGGTCTCAAT----- 2859  
Qy 510 ArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAla 529  
Db 2860 -----CCTACTGGGGGTTTCAGGA----- 2877  
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549  
Db 2878 -----CCTACTGGGGTCAATCCCATG----- 2899  
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566  
Db 2899 -----GGTGGCACACCAAGCTTGGGGTTCATCGATGCAAGGGCGGACCCCAACCATGG 2952  
Qy 567 -----ProIleHisGlyGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHis 584  
Db 2953 GGTCAACCCCTCAGGCCCAATCCAGGGGTACCTCGAGGTAGCAGTATAGGAGTCCGCAAT 3012  
Qy 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerAsnGlnArgSerProArg 601  
Db 3013 AGCCCCCAGGCTCTGAGGCGGACACTTCTGGGGACGACGACAGGCGCATGAGCACA 3072  
Qy 602 SerAsnSer 604  
Db 3073 AGCAGCAGT 3081

Search completed: February 28, 2004, 03:15:01  
Job time : 618 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 28, 2004, 00:19:58 ; Search time 131 Seconds  
(without alignments)

2579.886 Million cell updates/sec

Title: US-09-890-475-1

Perfect score: 3104

Sequence: 1 MSNPPVPAQPTTANPLL.....RYLGLSNQSPRNSLDPK 609

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USPRO\_spool\_p/US09890475/runat\_24022004\_135312\_5317/app\_query.fasta\_1.775  
-DB=Issued Patents NA -QFWT=frascap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09890475 @CGN 1 1 103 @runat\_24022004\_135312\_5317 -NCPU=6 -ICPU=3  
-NM MAP -LARGESQURY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/6CTUS COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	4.9	3312	3	US-09-060-410-1
2	151	4.9	3312	4	US-09-723-458-1
3	148.5	4.8	3824	4	US-09-688-188B-20
4	148.5	4.8	3824	4	US-09-291-417D-20
5	137.5	4.4	6558	4	US-09-491-356C-7
6	133	4.3	17612	3	US-08-911-853-29
7	133	4.3	17612	3	US-09-479-409-29
8	133	4.3	17612	4	US-09-479-453-29
9	131	4.2	9775	3	US-08-977-171-1
10	130.5	4.2	7393	4	US-09-620-312D-372
11	129	4.2	3324	4	US-09-620-312D-1020
12	127.5	4.1	3605	3	US-09-098-901-1

13	127.5	4.1	10432	4	US-09-919-172-97	Sequence 97, Appl
14	127.5	4.1	10432	4	US-09-978-594-21	Sequence 21, Appl
15	127	4.1	6306	1	US-08-466-390-3	Sequence 3, Appl
16	127	4.1	6306	1	US-08-470-950-3	Sequence 3, Appl
17	127	4.1	6306	1	US-08-467-781-3	Sequence 3, Appl
18	127	4.1	6306	1	US-08-195-487-3	Sequence 3, Appl
19	127	4.1	6306	2	US-08-483-924-3	Sequence 3, Appl
20	127	4.1	6306	5	PCT-US93-06160-3	Sequence 3, Appl
21	126.5	4.1	9934	3	US-08-977-171-2	Sequence 2, Appl
22	122.5	3.9	7100	4	US-09-308-375-1	Sequence 1, Appl
23	121.5	3.9	6436	4	US-09-600-099-1	Sequence 1, Appl
24	120	3.9	11220	3	US-09-105-537-32	Sequence 32, Appl
25	120	3.9	36778	3	US-09-105-537-5	Sequence 5, Appl
26	119.5	3.8	5053	2	US-08-685-576-2	Sequence 2, Appl
27	118	3.8	38506	4	US-09-320-878-19	Sequence 19, Appl
28	118	3.8	38506	4	US-09-141-908-1	Sequence 1, Appl
29	118	3.8	38506	4	US-09-657-440-19	Sequence 19, Appl
30	117.5	3.8	2487	1	US-08-261-304-1	Sequence 1, Appl
31	116.5	3.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl
32	116.5	3.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
33	116	3.7	1512	3	US-08-911-853-6	Sequence 6, Appl
34	116	3.7	1512	3	US-09-479-409-6	Sequence 6, Appl
35	116	3.7	1512	4	US-09-479-453-6	Sequence 6, Appl
36	116	3.7	2680	2	US-08-533-306A-5	Sequence 5, Appl
37	116	3.7	2680	2	US-08-743-923A-5	Sequence 5, Appl
38	116	3.7	2887	2	US-08-533-306A-3	Sequence 3, Appl
39	116	3.7	2887	2	US-08-743-923A-3	Sequence 3, Appl
40	116	3.7	7453	4	US-09-620-312D-248	Sequence 248, App
41	116	3.7	7501	1	US-09-620-312D-249	Sequence 249, App
42	116	3.7	10136	1	US-08-353-700-2	Sequence 2, Appl
43	116	3.7	10136	5	PCT-US95-16216-2	Sequence 2, Appl
44	115.5	3.7	2859	4	US-09-252-991A-944	Sequence 944, App
45	115.5	3.7	9626	4	US-09-150-867-2	Sequence 2, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-060-410-1  
; Sequence 1, Application US/09060410  
; Patent No. 6165461  
; GENERAL INFORMATION:  
; APPLICANT: Cobb, Melanie  
; APPLICANT: Hutchinson, Michele  
; APPLICANT: Chen, Zhu  
; APPLICANT: Berman, Kevin  
; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE  
; TITLE OF INVENTION: THEREFOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/060,410  
; FILING DATE: 14-APR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 860098.421  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 3312 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 121..3123  
 US-09-060-410-1

Alignment Scores:  
 Pred. No.: 7,19e-05 Length: 3312  
 Score: 151.00 Matches: 150  
 Percent Similarity: 34.75% Conservativity: 95  
 Best Local Similarity: 21.28% Mismatches: 237  
 Query Match: 4.86% Indels: 223  
 DB: 3 Gaps: 32

US-09-890-475-1 (1-609) x US-09-060-410-1 (1-3312)

QY 3 AsnTyr-----ProThrValAlaAlaGlnProThrThrThraAlaAsn 17  
 DB 1333 AATTACCAAGAAAGGAGATCTTAGAACAGAGCATCAGCTCCACAGCTCTCCA----- 1386  
 QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysLeuValGlu 37  
 DB 1387 CTTCAAGTCTCTGTCACAAATCATTATCGTAATAGAACACCTTTGCACTATACGA 1446  
 QY 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysLeuProGlnPheLeuLys 57  
 DB 1447 ACAGCATCATCTG-----GTTACAAGACAGATCAAGAACATGAG 1485  
 QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77  
 DB 1486 CAGGACTCTGAACCTTAGAACACAGATGCTGTTTAAAGCGGATGAGCGACAGCAT--- 1542  
 QY 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97  
 DB 1543 -----CAGAAGCAGCTGATGACTCTGGAAATTAACCTGAAGGCAGAAATGGACGAACAT 1596  
 QY 98 GlyVal-----ValleuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113  
 DB 1597 CGGCTCAGATTAGACAAAGATCTTGAACCTCAGCGCAACATTTCTGTCAGAAATG--- 1653  
 QY 114 SerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133  
 DB 1654 -----GAGAACTATTAAAGAACACCAACAGCTTCTATG 1686  
 QY 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153  
 DB 1687 GAAAAAGAGGCTAAAGTGATGCCCAACGAGGAG----- 1719  
 QY 154 SerLysGlyLeuArgLysIleThrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173  
 DB 1720 -----AAAAATTCACACACATTT-----CAGCTCAACAGAGAAA 1758  
 QY 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysIle 193  
 DB 1759 GAACCTGAATAGCTTTTGGAGCTCAAAAAAGAGAATATAAACTTCGAAAAGAGCAGCTT 1818  
 QY 193 eGlyLysPheThrLeuGlnGlyArgAlaPheThrLysGluSer-----ProMetSe 211  
 DB 1819 AAGGAGGAGCTGAATGAACACAGAGCACACCTTAAAGAAAGAAAGAGGAATGGCTTTCA 1878  
 QY 211 rSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGlu 231  
 DB 1879 AAGCAGAAGGAGATATTCAACATTTTCAGGCAGAGAGAGAGCTAATCTTCTCGACGT 1938  
 QY 231 yLysGlyLysVal-----LysIleGluSerTrpIleLysAspGluAlaGluTh 247  
 DB 1939 CAAGGACGATCTAGAGCTAGAATGTCGTCCTTCA-----AAAGAGATGTTACTT 1992  
 QY 247 zAlaAlaValAlaTrpArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267

DB 1993 GGTGGCATAAAGCTTGAACAGAGCCTTG-----TCAGGAGAGGTT-----AAACAAA 2040  
 QY 267 yMetAspAlaArgGlyLeuLeu-----LeuValAlaC 279  
 DB 2041 AGCAGACTCAGAGGACTTAGAACATGCAATGTTACTCGCAGACAGCATGAATCCATGCAA 2100  
 QY 279 yPheGlyVal---ProSerAsnPheArgSerThrAspLeuLeu-AspLeuLeuArgMet 297  
 DB 2101 GAACCTGGAGTTTCGCCACCTCAACATATTTCAGAAAGATCGCTGTGAGTTGATCAGACTG 2160  
 QY 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317  
 DB 2161 CAACATCAAACTGAGCTTACTACACAGCTGGAAATCAATAAGAGAAGG----- 2208  
 QY 318 SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal 337  
 DB 2209 -----GAACGGGAACCTAAGACGG---AAACATGTC----- 2235  
 QY 338 TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357  
 DB 2236 -----ATGAAGTTCGACAGCAGCCTTAAGAGTTTGAAGTCTAAAGAACTCCAA 2283  
 QY 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375  
 DB 2284 ATAAAGAGCAGTTTCAGGATACCTGCAAAATTCAAACCCAGACAGTACAAAGCATTAAGG 2343  
 QY 376 -----PheLys 377  
 DB 2344 AATCAGCTACTGAGACTACACAAAGAGTGAGCACAAAGCTGTTCTGAAAAGACTCAAG 2403  
 QY 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392  
 DB 2404 GAGAACAGCTCGGAAGTTAGCCATCTTGGCTGACGACATGATCATAGCATTATAGAA 2463  
 QY 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu----- 404  
 DB 2464 ATGCTCTCCACACAAGCTGCTGCTTTGGATGAAGCAGACAGAGAGCAGATGCCAGGTTTG 2523  
 QY 405 -----LeuProGlyTrpGlnIleLys----- 411  
 DB 2524 AAGATGAGCTACAGCAGGAACTGGAGCTGTTGATGCATATCAGAGCAAAATCAAGATG 2583  
 QY 412 -----GluGlnIleValSerLeu----- 417  
 DB 2584 CAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTGGAACAAAGGCTCTCCCTCGG 2643  
 QY 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429  
 DB 2644 AGAGCACTCTTAGAACAGAGATTGAAGAAGAGATGTTGGCTTTGCAGATGAACGACA 2703  
 QY 430 GluLysAlaArgSerLeuSerLeuMetGluAlaAlaLeuAlaLysArgMetTyrAsn 449  
 DB 2704 GAACGATAGCT-----AGCTGCTCGAGCGCCAGCCAGAGAAATTCAGAGCTTTGAC 2757  
 QY 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSer 469  
 DB 2758 TCTGAAAGCATGAGATAGGTTTGTAGTAACTAGTCTCTT-----TCTAAT 2802  
 QY 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489  
 DB 2803 CTCTCCCTCAGGATTCAGCCACAGCTACCCAGGA----- 2838  
 QY 490 IleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509  
 DB 2839 -----GCTTCTAGCTGGTCTCACAAT----- 2859  
 QY 510 ArgSerProGluTyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAla 529  
 DB 2860 -----CCTACTGGGGGTTTCAGAA----- 2877  
 QY 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyArgLeuHisArgGln 549

Db 2878 -----CCTCACTGGGGTCAATCCCATG----- 2898  
Qy 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566  
Db 2899 -----GCTGGCACACCAAGCTTGGGTCAATCCGATGCAAGCGGACCCCAACCATGG 2952  
Qy 567 -----ProHisGlyGlnGlnLeuProTyrGlyGlnArgValTyrArgHis 584  
Db 2953 GGTCAACCTCAGGCGCAATCAAGGGTACCTCGAGTAGCATATAGAGTCCGCAAT 3012  
Qy 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerArgGlnA-gSerProArg 601  
Db 3013 AGCCCCAGGCTCTGAGCGGACAGCTTCTGGGGACGGACGGAACAGGGCATGAGCAGA 3072  
Qy 602 SerAsnSer 604  
Db 3073 AGCAGCAGT 3081

RESULT 2  
US-09-723-458-1  
; Sequence 1, Application US/09723458  
; Patent No. 6586242  
; GENERAL INFORMATION:  
; APPLICANT: Cobb, Melanie  
; Hutchinson, Michele  
; Berman, Kevin  
; Chen, Zhu  
; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE  
; THEREFOR  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/723,458  
FILING DATE: 27-No. 6586242-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/060,410  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 860098,421  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3312 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 121...3123  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-723-458-1

Alignment Scores:  
Pred. No.: 7,19e-05 Length: 3312  
Score: 151.00 Matches: 150  
Percent Similarity: 34.75% Conservative: 95  
Best Local Similarity: 21.28% Mismatches: 237

Query Match: 4.86t Indels: 223  
DB: 4 Gaps: 32  
US-09-890-475-1 (1-609) x US-09-723-458-1 (1-3312)  
Qy 3 AsnTyr-----ProProThrValAlaAlaGlnProThrThrThraAlaAsn 17  
Db 1333 AATTACCAAGAAAGGAGATCTTAGAACAAGAGCATCAGCTCCACAGTCTCCA----- 1386  
Qy 18 ProLeuLeuGlnArgHisGlnSerGlnArgArgGluLeuProLysIleValGlu 37  
Db 1387 CCTCAAGTCTCGTCACAAATCATTATCGTAATAGAGACACTTTCGAACATATACGA 1446  
Qy 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57  
Db 1447 ACAGCATCACTG-----GTTACAAGACAGATGCAAGAACATGAG 1485  
Qy 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77  
Db 1486 CAGGACTCTGAACCTTAGAGAACAGATGCTGTGTTATAAGCGGATGAGCGCAGCAT--- 1542  
Qy 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97  
Db 1543 -----CAGACAGCTGATGACTCTGGAATAAATACTGAAGCGCAGAATGACCAACAT 1596  
Qy 98 GlyVal-----ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeu 113  
Db 1597 CGGCTCAGATTAGACAAAGATCTTGAAACTCAGCGCAACAATTCGCTGCAGAAATG--- 1653  
Qy 114 SerProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSerGln 133  
Db 1654 -----GAGAACTTATTAAAGAAACACCAAGCTTCTATG 1686  
Qy 134 GluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153  
Db 1687 GAAAGAGAGCTAAAGTGATGCCAACAGGAG----- 1719  
Qy 154 SerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGlu 173  
Db 1720 ---AAAAAATTCACACACACATT-----CAGGCTCAACAGAGAAA 1758  
Qy 174 GluIleProSerAlaLeuLys-LeuAlaLysGluProAlaLysPheValLeuAspCysI1 193  
Db 1759 GAACGTAATAGCTTTTGGAGTCTCAAAAAGAGATATAAACTTCGAAAAGAGCAGCTT 1818  
Qy 193 eGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSer-----ProMetSe 211  
Db 1819 AAGGAGGAGCTGAATGAAACCCAGAGCACACCTAAAAAAGAAAGCAGGAATGGCTTTCA 1878  
Qy 211 rSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgI 231  
Db 1879 AAGCAGAAGGAGAAATATTCAACATTTTCAGGCAGAGAAGAAGCTTAATCTTTCGACCT 1938  
Qy 231 YLysGlyLysVal-----LysIleGluSerTrpIleLysAspGluAlaGluTh 247  
Db 1939 CAAAGGCAGTATCTAGAGCTAGATGTCGTCGCTTCA-----AAAGAAGATGTTACTT 1992  
Qy 247 rAlaAlaValAlaTTPArgLysArgLeuMetThrGlu-GlyGlyLeuAlaAlaGluL 267  
Db 1993 GGTCCGCAATAACTTGAACAGGACCTTG-----TCAGGGAGGAGT-----AAACAA 2040  
Qy 267 ysMetAspAlaArgGlyLeuLeuLeu-----LeuValAlaLac 279  
Db 2041 AGGCAGACTCAGAAGGACTTAGAACATGCAATGTTTACTCGCAGACGATCAATCCATCAA 2100  
Qy 279 ysPheGlyVal----ProSerAsnPheArgSerThrAspLeuLeu-AspLeuIleArgMet 297  
Db 2101 GAACCTGGAGTTTCGCCACCTCAACACTATTTCAGAGATCGCTGTGAGTGTGATCAGACTG 2160  
Qy 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317  
Db 2161 CAACATCAAACTGAGCTTACTTAACACGCTGGATACATACAAATAGAGAAGG----- 2208

QY 318 SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal 337  
Db 2209 -----GAACGGAACTAAGACGG-----AAACATGTC----- 2235  
QY 338 TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357  
Db 2236 -----ATGGAAGTTTCAGACAGCCTAAGAGTTTGAAGTCTAAAGAACTCA 2283  
QY 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375  
Db 2284 AFAAAAAGCAGTTTCAGGATACCTGCAAAATTCAAACACAGACAGTACAAAGCATTAGG 2343  
QY 376 -----PheLys 377  
Db 2344 AATCACTAGTGGAGACTACCAAGAGTGGACCAAAAGCTGTTCTGAAAGACTCAAG 2403  
QY 378 GluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetGln 392  
Db 2404 GAGGAACAGACTCGGAAGTAGTCCATCTTGGCTGAGCAGTAGTATCATAGCATTAATGAA 2463  
QY 393 CysMetGluThrHis-----LysLeuAspProAlaLysGlu----- 404  
Db 2464 ATGCTCTCCACAAGCTCTGCGTTTGGATGAAGCACAGGAGCAGAAATGCCAGGTTTGT 2523  
QY 405 -----LeuProGlyTyrGlnIleLys----- 411  
Db 2524 AAGATCCAGCTACAGCAGAACTGGAGCTGTTGTAATGCATATCAGACAAATCAAGATG 2583  
QY 412 -----GluGlnIleValSerLeu----- 417  
Db 2584 CAGGCTGAGGCCCAATGATCGAGCTTCGAGAGCTGGAAACAAAGGCTCTCCCTCGG 2643  
QY 418 -----GluLysAspThrLeuGlnLeuAspLysGluMetGlu 429  
Db 2644 AGAGCACTCTTAGACAGAGATTGAAGAGAGAGTTGGCTTTCAGAGTGAAGCCACA 2703  
QY 430 GluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsn 449  
Db 2704 GAACGAATACGT-----AGCCTGCTCGAGCCCGCCAGGCGAGAGAAATTCAGCTTTTAC 2757  
QY 450 GlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSer 469  
Db 2758 TCTGAAGCATGAGATTAGTTTGTATACATGGTCCCTT-----TCTAAT 2802  
QY 470 TyrSerProIleTyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGlu 489  
Db 2803 CTCTCCCTGAGCATTCAGCCACAGCTACCCAGGA----- 2838  
QY 490 IleSerAlaLeuValSerTyrLeuGlyProSerThrSerPheProHisArgSerArg 509  
Db 2839 -----GCTTCAGTGGTCTCAAT----- 2859  
QY 510 ArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyrAla 529  
Db 2860 -----CCTACTGGGGGTTCAGGA----- 2877  
QY 530 TyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGln 549  
Db 2878 -----CCTACTGGGGGTTCATCCCATG----- 2898  
QY 550 TyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerPro----- 566  
Db 2899 -----GGTGGCACACCAAGCTGGGTTCATCCGATCAAGGGGACCCCAACCATGG 2952  
QY 567 -----ProIleHisGlyGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHis 584  
Db 2953 GGTACCCCTCAGGGCCCAATGCAAGGGGTACCTCGAGGTAGGATATAGAGTCCGCAAT 3012  
QY 585 SerProSerGluGluArgTyrLeu-----GlyLeuSerAsnGlnArgSerProArg 601  
Db 3013 AGCCCCAGGCTCTGAGCGGACAGCTTCTGGGGGACGCGACGGAACAGGGCATGAGCAGA 3072  
QY 602 SerAsnSer 604

Db 3073 AGCAGAGT 3081  
RESULT 3  
US-09-688-188B-20  
: Sequence 20, Application US/09688188B  
: Patent No. 6656716  
: GENERAL INFORMATION:  
: APPLICANT: FLOWMAN, GREGORY  
: APPLICANT: MARTINEZ, RICARDO  
: APPLICANT: WHYTE, DAVID  
: TITLE OF INVENTION: STB20-RELATED PROTEIN KINASES  
: FILE REFERENCE: 038602/0328  
: CURRENT APPLICATION NUMBER: US/09/688,188B  
: CURRENT FILING DATE: 2000-10-16  
: PRIOR APPLICATION NUMBER: 09/291,417  
: PRIOR FILING DATE: 1999-04-14  
: PRIOR APPLICATION NUMBER: 60/081,784  
: PRIOR FILING DATE: 1998-04-14  
: NUMBER OF SEQ ID NOS: 155  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 20  
: LENGTH: 3824  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-09-688-188B-20  
Alignment Scores:  
Pred No.: 0.000164 Length: 3824  
Score: 148.50 Matches: 150  
Percent Similarity: 34.88% Conservative: 98  
Best Local Similarity: 21.10% Mismatches: 227  
Query Match: 4.78% Indels: 237  
DB: 4 Gaps: 35  
US-09-890-475-1 (1-609) x US-09-688-188B-20 (1-3824)  
QY 3 AsnTyr-----ProProThrValAlaAlaGlnProThrThrAlaAsn 17  
Db 569 AATTACAGAGAAGAGGAGATCCTGACAAAGCATCATCCAACTCCA----- 622  
QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37  
Db 623 CCCCAAGTATCTCGTCACAAATCACACTATCGTAATCGAAGACACTTGTCTACTATACGG 682  
QY 38 ThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPheLeuLys 57  
Db 683 ACAGCATCACTG-----GTTACGAGGCAATGCAAGAACATGAG 721  
QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77  
Db 722 CAGGACTCTGAGCTTAGAGAACAAATCTCTGGCTATAGCGAATGAGGCGACAACTAT--- 778  
QY 78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97  
Db 779 -----CAAAAGCAACTGATGACTCTGGAATAAAGCTAAAGCTGAGTGGATGAACAT 832  
QY 98 GlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProProArg 117  
Db 833 CGCTCTCAGATTA-----GACAAAGATCTTGAACATCAGCGT 868  
QY 118 AsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValPro 137  
Db 869 AACAAATTTGCTGCAGAAATGAGAACTTATCAAGAAACACAGGCTGCGATGGAGAAA 928  
QY 138 GluThr-----SerAsnLysProGluGlyGlyArgMetCysGluLeuMetCysSer 154  
Db 929 GAGCTAAAGTATGTCCTCAATGAA-----GAG 955  
QY 155 LysGlyLeuArgTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGlu 174  
Db 956 AAAAAATTCAGCAACATATT-----CAGGCCCAACAGAAAGAAA 997





623 CCCAAGTATCTCGTCACAAATCACTATCGTAAATCGAGAACACTTTGCTACTATACGG 682  
Db  
38 ThrGluSerThrSerMetAspIleThrIleGlnSerLysGlnProGlnPheLeuLys 57  
Qy  
683 ACAGCATCACTG-----GTTACGAGGCAAAATCGAAGACATGAG 721  
Db  
58 SerIleAspGluLeuAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77  
Qy  
722 CAGGACTCTGAGCTTAGAGAACAAATGCTCGGTATTAAGCGAATGAGGCGGACACAT--- 778  
Db  
78 AspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsn 97  
Qy  
779 -----CAAAAGCACTGATGCTGGAACCAAGCTAAAGGCTGAGATGATGAACAT 832  
Db  
98 GlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSerProArg 117  
Qy  
833 CGCCTCAGATTA-----GACAAAGATCTTGAACATCGACGT 868  
Db  
118 AsnAsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValPro 137  
Qy  
869 AACAAATTTGCTGCAAGAAATGAGAAACTTATCAAGAAACACAGGCTGCCATGAGAAA 928  
Db  
138 GluThr-----SerAsnLysProGluGlyArgMetCysGluLeuMetCysSer 154  
Qy  
929 GAGGCTAAAGTATGATCCCAATGAA-----GAG 955  
Db  
155 LysGlyLeuArgLysIleThrAlaAsnIleSerAspGlnAlaLysLeuMetGluGlu 174  
Qy  
956 AAAAAATTCACACATATT-----CAGGCCACACAGAAAGAA 997  
Db  
175 IleProSerAlaLeu-----LysLeuAlaLysGluProAlaLys 187  
Qy  
998 CTGAATAGTTTCTCGAGTCCAGAAAAGAGAGATATAAATCTCGAAAAGACGACTTAA 1057  
Db  
188 PheValLeuAspCysIleGlyLysPheThrLeuGlnGlyArgArg-----AlaPhe 204  
Qy  
1058 GAGGAGCTAAAT---GAAACCCAGAG-TACCCCCAAAAGAAAAGACAGGAGTGCCTTC 1113  
Db  
205 ThrLysGluSer-----ProMetSerSer 212  
Qy  
1114 AAAGCAGAGGAGAGATATACACATTTCCAGCAGAGAGAGAGTAACTTCTTCGACG 1173  
Db  
213 AlaArgGlnValSerLeuLeuLeuLeuLeuLeuLeuLeuMetProAspArgGlyLys 232  
Qy  
1174 TCAAGACAAATA-----CCTAGAGCTGGAATG 1200  
Db  
233 GlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTrp 252  
Qy  
1201 CCGTCGCTTCAAGAGAGAT-----GTTACTTGG 1230  
Db  
253 ArgLysArgLeuMetThr-----GluGlyGlyLeuAlaAlaGluLysMetAsp 269  
Qy  
1231 GCGTCATAACTTAGAGCAGGACCTTGTTCAGGAGGAGT-----AAACAAAAGACAGAC 1284  
Db  
270 AlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer----- 284  
Qy  
1285 TCAGAAGGACTTAGA-----GCATGCCATGTCTACTCCGACAGCATGATCTATGCA 1335  
Db  
285 -----AsnPheArgSerThrAspLeuLeu-AspLeuIleArgMe 297  
Qy  
1336 AGAAGCTGGAGTTCCGCCACCTCAACACAAATTCAGAAGATGCGCTGTGATGATCAGATT 1395  
Db  
297 tSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVa 317  
Qy  
1396 ACAGCATCAAACTGAGCTCACTAACCGCTGGAATATATAAGCGAAGA----- 1444  
Db  
317 lSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVa 337  
Qy  
1445 -----GAACGAGAACTAAGACGA---AAGCATGTC----- 1471  
Db  
337 lTyThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357  
Qy  
1472 -----ATGGAAGTTTCGACCAACAGCGCTTAAGATTTGAAGTCTTAAAGAACTCCA 1518  
Db

357 sMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAla----- 375  
Qy  
1519 AATAAAAAGCAGATTTCAGGATACCTGCAAAATCCAAACCCAGACAGTAGTCAAAAGCATTAAG 1578  
Db  
376 -----PheLys 377  
Qy  
1579 AAATCACCTGCTGGAGACTACACCAAGAGTGCAGACAAAGCTGTTCTGAAACGGCTCAA 1638  
Db  
377 sGluAlaAlaThrLysGlnLeuAlaValLeuSer-----SerValMetG1 392  
Qy  
1639 GGAGGAACAGACCCCGAAATTAGCTATCTTGCTGAGCAGTATGATCAAGCATTAATGA 1698  
Db  
392 nCysMetGluThrHis-----LysLeuAspProAlaLysGluLeuProGlyTrpGlnI1 410  
Qy  
1699 AATGCTCTCCACACAGCCCTGCGTTTGATGAACACAGAGGACGAGTGCAGGTTT 1758  
Db  
410 eLysGluGlnIle-----ValSerLe 417  
Qy  
1759 GAAGATGACGCTGCAGCAGAACTGAGCTGTTGAATGCGTATCAGAGCAAAATCAAGAT 1818  
Db  
417 uGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLe 437  
Qy  
1819 GCAGCTGAGGCA---CAACATGATCGAGAGCTTCGCGAGCTTGAACAGAGGTTCTCCT 1875  
Db  
437 u-----MetGluGluAlaAlaLeuAlaLysArgMetTyra 449  
Qy  
1876 CCGGAGGCGACTCTTTAGAAACAAAGATTGAAGAGAGATGTTGGCT-----TTGCAGAA 1929  
Db  
449 nGlnGlnIleLysArgProArgLeuSerProMetGluMetProProValThrSerSerSe 469  
Qy  
1930 TGAGCCACAGAACGAATACGA----- 1951  
Db  
469 rTySerProIleTyArgAspArgSerPheProSerGlnArgAspAspGlnAspG1 489  
Qy  
1952 -----AGCCTGTTTGAACGCTCAAGCCACAGAGA 1977  
Db  
489 uIleSerAlaLeuValSerSerTyrrLeuGlyProSerThrSerPheProHisArgSerAr 509  
Qy  
1978 GATTGAAGCTTTGACTCTGAAAGCATGAGACTAGGTTTTAGTAATATATGCTCTTCTTAA 2037  
Db  
509 gArgSerProGluTyrrMetValProLeuProHisGlyGlyLeuGlyArgSerValTyra1 529  
Qy  
2038 TCTCTCCCTCGAG-----GC 2052  
Db  
529 aTyrrGluHisLeuAlaProAsn-----SerTyrrSerProGlyHisGly----- 543  
Qy  
2053 ATTCAAGCCACAGCTACCCGAGAGCTTCTGTTGGTCACACAAACCTTCTGGGGGTCCAGG 2112  
Db  
544 ---HisArgLeuHisArgGlnTyrrSerProSerLeuValHisGlyGlnArgHisProLe 562  
Qy  
2113 ACCTCACTGGGTCATCCCATGGTGGCCACCAACCAAGCTTGGGCG-----CATCCAT 2166  
Db  
562 uGlnTyrrSerPro-----ProIleHisGlyGlnGlnGlnLeuProTyrrG1 577  
Qy  
2167 GCAAGGTGGACCCAGCCATGCGGCTCACCTTTCAGGGCCAATGCAAGGGGTACCTCGAG 2226  
Db  
577 yIleGlnArgValTyrrArgHisSerProSerGluGluArgTyrrLeu-----GlyLe 594  
Qy  
2227 TAGCAGTATGGGATCGCGCAATAGCCCCCAGGCTCTGAGCGGACAGCTTCTGGGGGACG 2286  
Db  
594 uSerAsnGlnArgSerProArgSerAsnSer 604  
Qy  
2287 GACGAGCAGGGCATGAGCAGAGACGAGT 2317  
Db

## RESULT 5

US-09-491-356C-7

; Sequence 7, Application US/09491356C

; Patent No. 6566061

; GENERAL INFORMATION:

; APPLICANT: Philibert, Robert A.

; APPLICANT: Gibbs, Edward I.

; APPLICANT: Delisi, Lynn

1 TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF X013

2 FILE REFERENCE: 9465.6US11  
3 CURRENT APPLICATION NUMBER: US/09/491.356C  
4 PRIOR FILING DATE: 2000-01-26  
5 PRIOR APPLICATION NUMBER: PCT/US99/09365  
6 PRIOR FILING DATE: 1999-04-29  
7 PRIOR APPLICATION NUMBER: 60/083,465  
8 PRIOR FILING DATE: 1998-04-29  
9 NUMBER OF SEQ ID NOS: 24  
10 SOFTWARE: PatentIn version 3.1  
11 SEQ ID NO 7  
12 LENGTH: 6558  
13 TYPE: DNA  
14 ORGANISM: Mus musculus  
15 US-09-491-356C-7

## Alignment Scores:

Pred. No.:	0.00535	Length:	6558
Score:	137.50	Matches:	137
Percent Similarity:	33.46%	Conservative:	118
Best Local Similarity:	17.98%	Mismatches:	226
Query Match:	4.43%	Indels:	281
DB:	4	Gaps:	36

US-09-890-475-1 (1-609) x US-09-491-356C-7 (1-6558)

Qy	17	AsnProLeuLeuGlnArgHisGlnSerGluGlnArgargGluLeuProLysIleVal	36
Db	3928	GACCAGTGTTCAGTAGTGCACGAGCCGAGCCG	3975
Qy	37	GluThrGluSerThrSerMetAspIleThrIleGlnSerLysGlnProGln	54
Db	3976	TGCTACCCACATCGACTGCTGACATCAGGATGAGAGAAACCCGAGCGCGCAT	4035
Qy	55	-----PheLeuLysSerIleAspGluLeuAlaPheSerValAlaValGluThrPhe	72
Db	4036	AAACGTATTTCTCAAGATTTTAGACCAATGG	4071
Qy	73	LysArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSer	92
Db	4072	CGCAGTCTCTTTGGAGTACAGTGTGATGATCAGCAGACCCCAACACT	4122
Qy	93	LysLeuGluSerAsnGlyValValLeuAlaAlaAlaAlaAlaAlaAlaAlaAla	112
Db	4123	-----GAGATGAATCTCTCTTG	4140
Qy	113	LeuSerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSer	132
Db	4141	-----GAGAACATTTGCC-----AAGGCCAATTCGAGGTTTCCACAGTCA	4182
Qy	133	GlnGluIleValProGluThr-----SerAsnLysProGluGlyGlyArgMet	148
Db	4183	GCAGAGCGGGTCTATCTCTGGAAGTACAGACAGCAACATGCGCCAGCAGCAGCACC	4242
Qy	149	CysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGln	168
Db	4243	AAACCTGTGCTCAGCTCT-----CTAGAGCGA-----TCTGTGTATGGTTGGTG	4287
Qy	169	AlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPhe	188
Db	4288	GCTCTCTCATGTGCAAACTGCCACITTCAGTCCAG-----GGCCAT	4329
Qy	189	ValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg-----AlaPheThr	205
Db	4330	GTATTAAGAGTGTGGGAGAGAACTGGAAAGGGTCAGCACCTGGGCTCTTCTCCCGT	4389
Qy	206	LysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu	225
Db	4390	AAAGAACGAGATCGACAAACAGAGAGATGTCCTGTTGAGCCCAACAGCCTTTCTTA	4449
Qy	226	-----LeuMetProAspArgLysGlyLysValLysIleGluSerTrp	240
Db	4450	TCACTGTGTCTAAACATGTCTGAAAGGACAGGATGAGCAGCGCGGAGGACTCTGGCCTCC	4509

Qy	241	IleLysAspGluAlaGluThrAlaAlaValAlaTrpArg-----	253
Db	4510	CTCCACAGCCAGGTGCACAGATTGTGATTAATTGGCGAGAAACACAGTACTTAGATGAT	4569
Qy	254	-----LysArgLeuMetThrGlu-----GlyGly	261
Db	4570	TGCAAAACCAAGCAGCTAATCATGATGAGGCATCAAACTCGGCTCAACTGGTGGGGGC	4629
Qy	262	Leu-----AlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeu	275
Db	4630	ATGTTTGACACTGTGCAGGTAGTACCCAGCAGACTACAGAGTGGCGCCAGCTTCTCCT	4689
Qy	276	LeuValAlaCysPheGlyVal-----ProSerAsnPheArgSerThrAspLeu	291
Db	4690	GAGATCATCATCAGCGGCACCTGTGCAGTGTAAATGAGCTCTTACTACTGTG	4749
Qy	292	LeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGln	311
Db	4750	TTGGACATGCTGAGCGTCTTATCAACGGAACGTTGGCTGCAGACATG-----	4797
Qy	312	PheLeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMethIle	331
Db	4798	-----TCCAGTATCTCGCAAGGC-----	4815
Qy	332	GluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuVal	351
Db	4816	-----AGCATGGAGGAAACAAACGCTGCATATATGAAC	4848
Qy	352	LeuThrSerPheLeu-----LysMetSerLysGluSerPheGluArgAla	366
Db	4849	CTGGTGAAGAAGCTTCAGAAAGACTTGGGGGAGCGCCCAATCAGACAGTCTGAGAAGGT	4908
Qy	367	LysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThr-----	381
Db	4909	CACCAACTGTTCCACTACCCAGCAGAACCCAGATGTCAACTGTGTGAGCCACAGGCG	4968
Qy	382	-----LysGlnLeuAlaValLeuSerSerVal-----	390
Db	4969	TCCTTTATTCACCAAGGGCAACAAGATTGTGGCTTCGATTCCTTCAAGAAGGAG	5028
Qy	391	---MetGlnCysMetGluThrHisLysLeuAspProAlaLysGluLeuProLysTrpGln	409
Db	5029	GGTCTACAGTTTCTACCAACAAAGATCTCTCCC-----TGGGAG	5070
Qy	410	IleLysGluGlnIleValSerLeuGluLys-----AspThrLeu	422
Db	5071	CTTTTGTAGGGCTGAGCCATCAACAGCACCCTGTCATGGCGCTGTTTGGCAGATC	5130
Qy	423	GlnLeuAspLysGluMet-----GluGluLysAlaArgSerLeuSerLeuMetGlu	439
Db	5131	CGAGTGCACCGAGAGTGGCAGGGGGAGGAGCAGCGGCTGTG-----	5178
Qy	440	GluAlaAlaLeuAlaLysArgMetTyrAsnGlnIleLys-----ArgProArg	456
Db	5179	-----CTCTATCATACCCACTGAGCGCTCGACCCAGACCTAT	5217
Qy	457	---LeuSerProMetGluMetProValThrSerSerSerTyrSerProLysTrp	474
Db	5218	TACCTGGAACCACTACCTCTGCGCCCAAGAGATGAGGAGCCACAGCCCTCCCTACTA	5277
Qy	474	-----	474
Db	5278	GAGCTGAGAAAAAGGCTCCTGAGCCCCCAAGACTGACAAACAGGGGGCTCTCTCGG	5337
Qy	475	-----ArgAspArgSerPheProSerGln	482
Db	5338	AGCACTGAGGAGCGCAAAAAGAGTCTACCAAGGGCAAAAACGAGCCAGCCAGCCACC	5397
Qy	483	ArgAspAspAspGlnAspGluIleSerAlaLeuValSerTyrLeuGlyProSerThr	502
Db	5398	AAGAACGAGGACTATGGC-----ATGGGGCCAGGTCGG	5430

```
QY 503 SerPheProHisArgSerArgArgSerProGluTyrMet----- 515
|||
DB 5431 AGTGGCCCTATGTTGTGACAGTGCCTCCAGACCTTCTACACCATGCAAAATCTGGTTCT 5490
|||
QY 516 -----ValProLeu 518
DB 5491 ATATCCCACTTAGCTACAGGCAAGCTCCATGGGCTGTATATCCCAAAACAGCCACTA 5550
QY 519 ProHisGlyGly----- 522
DB 5551 CCTGCTGTTGGCCCTCGTGTGGATCCATACCGCCCGTGGCATACCAATGCAAAAGCTG 5610
QY 523 -----LeuGlyArgSerValTyrAlaTyrGluHisLeu 533
DB 5611 CCAACTCGACCAACTATCCCGGTGTGCTGCTCAACATATGCTACTGTCTATGGGCTA 5670
QY 534 AlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSer 553
DB 5671 GAACCTCTCTTTATAGACATCT-----GTATACCGGCGAGCAACCCACA 5718
QY 554 LeuValHisGlyGln-----ArgHisProLeuGlnTyrSerProProHisGlyGln 571
DB 5719 GTGCCCCAGGACAGCGCTTCGCCAACAGCTCCAGCAGAGTCAAGGGATGTTGGGACAG 5778
QY 572 -----GlnGlnLeuPro-----TyrGlyIleGlnArgValTyrArgHis 584
DB 5779 TCATCTGTCCATCATGATGACCCCTAGTTCTTCCTATGTTTGCAGACTTCCAGCTCTCT 5838
QY 585 SerProSerGluGluArgTyr-----LeuGlyLeuSerAsnGlnArg 598
DB 5839 TCTCTCTCTCCAGGGCTATACATCCATGTTCTTCATGTTGGGATTCAGCAACACACA 5898
QY 599 SerPro 600
DB 5899 GGCCCT 5904

RESULT 6
US-08-911-853-29
; Sequence 29, Application US/08911853
; Patent No. 6048710
; GENERAL INFORMATION:
; APPLICANT: Gerritse, Gijbert
; APPLICANT: Quax, Wilhelmus J.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED
; TITLE OF INVENTION: EXPRESSION LEVELS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-Seq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,853
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,092
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC361-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 17612 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-911-853-29
```

## Alignment Scores:

```
Pred. No.: 0.0803 Length: 17612
Score: 133.00 Matches: 131
Percent Similarity: 35.95% Conservative: 98
Best Local Similarity: 20.57% Mismatches: 224
Query Match: 4.28% Indels: 184
DB: 3 Gaps: 31
```

US-09-890-475-1 (1-609) x US-08-911-853-29 (1-17612)

```
QY 14 ThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuPro 33
DB 9754 ACCTCGATCCTCCCTCGACCAATCGTAGCTAGCCGCGCAACGCGCAACAGCGCGGTG 9813
QY 34 LysIleValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnPro 53
DB 9814 AAGATCGGCGAGATGAGCTGACCTCGCGCTTCATCGCC----- 9852
QY 54 GlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLys 73
DB 9853 -----TCGGCTTCGGCGCAACGCGGTGTTCTTCAAG 9885
QY 74 ArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLys 93
DB 9886 CAC-----CAGCGTCACGAAGACAAATAAGGTCATCCCTTCTCTGAACAG 9930
QY 94 LeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsn-----PheHisGln 110
DB 9931 CCCC GCCCGCTCGCGGGGCTTTTGTGTCACCCCTTACGTCCTCATCCTCTCGCCAG 9990
QY 111 ProMetLeuSerProProArgAsnValSerValGluThrThrValThrValSerGln 130
DB 9991 GCTGTGCTCGCGCT-----GCAAAATCGGCACCTGCAGTTTTGTGCG 10032
QY 131 ProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGlu 150
DB 10033 AAATCGTTAACTTGGCGCTCGGCCATGCCATAAAACAAACAGAACACAGCAAGATG 10092
QY 151 -LeuMetCysSerLysGlyLeuArgLysTyrIle-----TyrAlaAsnIleSerAspG 168
DB 10093 GATCTTCTGTTGGGGAACGATCGCCCATCTCCACGATACCCACGCGCCCTGACGG 10152
QY 168 nAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys-P 188
DB 10153 CTCCGCAAGCCCGCTTGGCGCGCTGCCCTTGCCTTCGCCAAACGCCACGCGGTGC 10212
QY 188 heValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluS 208
DB 10213 TGCTGCGGAGCCCTTCGGCCAGTCCAGTCCAGTGGCGCCGCT----- 10259
QY 208 erProMetSerSerAlaArgGlnValSerLeuIleLeuSerPheLeuLeuMetP 228
DB 10260 -----GCCAGCTCGCGCGCTGCAGGAG-----G 10284
QY 228 roAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluAla-----G 246
DB 10285 CCGAGCGCTTCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 10344
QY 246 luThrAlaAlaValAlaTrpArgLys-----ArgLeuMetThrGluGlyG 261
DB 10345 AGGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 10403
QY 261 lyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheG 281
DB 10404 -----ATGGGTGCCGAACCTTGACCTAGCCAGCGCTGCGCGCACTCTCC 10448
QY 281 lyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerA 301
```



D	b	9991	GCTGTGCTCCGCCCT-----GC AAAATCGGCAC TCGCAGT TTTTGCGC	10032
Q	y	131	ProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGlu	150
D	b	10033	AAATCGTTAACTGGCGCCTCGGCATCCATAAAAAACAAGAACAACAGCAAGATG	10092
Q	y	151	-LeuMetCysSerLysGlyLeuArgLysTrile-----TyrAlaenIlleSerAspG	168
D	b	10093	GATCTTCTCTTCGGGAACGCAATCCGCCCATGTCCACCGATACCACGCCCTCGTAGCGG	10152
Q	y	168	nAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys-P	188
D	b	10153	CTCCCGAAGCCCGCCTTCGCCCGCTGCCCTTCGCCTCGCCAAAACGCCACGCGCTGC	10212
Q	y	188	heValLeuAspCysIlleGlyLysPheTy-LeuGlnGlyArgArgAlaPheThrLysGluS	208
D	b	10213	TGCTGCGCAGGCCCTTCGGCCAGTTCACGTCCAGTCCAGTCCGCGCGGT-----	10259
Q	y	208	erPrometSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetP	228
D	b	10260	-----GCCAGCTGGCGCGCTGCAGAG-----G	10284
Q	y	228	roAspargGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluLa-----G	246
D	b	10285	CCCAGCGCTTCGCCGCCCGCTGCTGCCCTCACCTGGCTGGAGCCCGAGGCCTTCGAGC	10344
Q	y	246	IuThrAlaAlaValAlaTrpArgLys-----ArgLeuMetThrGluGlyG	261
D	b	10345	AGGAGCTGGCCCTTGGCCTACCGCGGCACTCTCCAGAGTGGCGAGATGCCGAGGCG-	10403
Q	y	261	lyLeuAlaAlaIaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPheG	281
D	b	10404	-----ATGGSTGGCGAAC TTGACTAGCCAGCCTGGCGGA ACTCACTCCC-----	10448
Q	y	281	lyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerA	301
D	b	10449	-----GAATCCGGCCACTCTCTGGAG-----	10469
Q	y	301	snglIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleV	321
D	b	10470	-----CAGGAAGATGACCGCCGATCATCCGCTGATCAACGCCATCC	10512
Q	y	321	algluSerSerIleLysArgGly-----MecHisIleGluAlaLeuGlu-----M	336
D	b	10513	TCAGCAGGCGATCAAGAGCGCGCGCTCCGACATCCACGTGAAACCTTCAGAAAACGCC	10572
Q	y	336	etValTyThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheL	356
D	b	10573	TGtGTGTGGCTTTCGCTCGACGGC-----ATCTCTCGCGAAGTGA	10614
Q	y	356	eulySmetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaP	376
D	b	10615	TCGAACCGCGCCCGAG-----CTGGCG-	10637
Q	y	376	helysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluT	396
D	b	10638	-----GCCTGCTGTCTCGGGTCAAGTCAAGTCAAGCG-	10670
Q	y	396	hrHisLysLeuAspProAla---LysGluLeuPro-----	406
D	b	10671	-----CGCTGGACATCCCGAGAGCGGTACCGCAGGAGCGCGTATTTCGCTCAAGG	10725
Q	y	407	----GlyTrpGlnIleLysGluGlnIleValSerLeu-----GluLysA	420
D	b	10726	TCGGCGTGCAGAGTGGATATCCGCTCTCCACCCTGCCGTCGCCCAACCGCGAGCGGG	10785
Q	y	420	spThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeu-----	437
D	b	10786	TGtGTGTGCTGTG-----CTCGACAGAGCGCGCGCGCTGTGCTCAGCATC	10836
Q	y	438	----MetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgPro-	455
D	b	10837	TGGGCATGACGAGCGC-----GACCGCGCTGTCTCGACACAACTGGGCAAGCGCG	10890

## RESULT 8

RESOL 8  
IIS-09-479-453-29

03-09-473-433-23 : Sequence 29. Application IIS/09479453

; Patent No. 6313283

; PALLELL NO. 6313283  
; GENERAL INFORMATION.

APPLICANT: Gerritse, Gishbert

APPLICANT: Quax Wilhelms J.

APPLICANT: Quax, Willemus J.  
TITLE OF INVENTION: EXPRESSION SYSTEM FOR ALTERED

;	TITLE OF INVENTION:	EXPRESSION SYSTEM
:	TITLE OF INVENTION:	EXPRESSION LEVELS
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

TITLE OF INVENTION:	EXE
NUMBER OF SENTENCES:	37

; NUMBER OF SEQUENCES: 3/  
; CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS: ADDRESS: Genecor International

;  
ADDRESS: GENENCOR INTERN  
STREET: 925 PACE MILL ROAD

STREET: 925  
CITY: DALTON, GA 30705

CITY: PALM  
STATE: CA

STATE: CA

; COUNTRY: USA  
ZTD: 043041013; ZIP: 94304-1013  
COMMITTEE REPLY FORM.

; COMPUTER READABLE FORM: MEDIUM TYPE: Diskette

; MEDIUM TYPE: Diskette

COMPUTER: IBM COMPATIBLE

```
; OPERATING SYSTEM: DOS
```

```

;
SOFTWARE: FastSEQ for

```

```

; CURRENT APPLICATION DATA:
;

```

APPLICATION ;

;  
FILING DATE: \_\_\_\_\_

;  
; PRIOR APPLICATION DATA:

**APPLICATION**

;  
FILING DATE: \_\_\_\_\_

ATTORNEY/AGENT INFORMATION:

; NAME: Glaister, Debra J

; REGISTRATION NUMBER: 33,888

REFERENCE/DOCKET NUMBER: GC

; TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-846-7620

; TELEFAX: 650-845-6504

; INFORMATION FOR SEQ ID NO:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 17612 base

TYPE: nucleic acid

```
TOPOLOGY: linear
US-09-479-453-29
Alignment Scores:
  Pred. No.: 0.0803 Length: 17612
  Score: 133.00 Matches: 131
  Percent Similarity: 35.95% Conservative: 98
  Best Local Similarity: 20.57% Mismatches: 224
  Query Match: 4.28% Indels: 184
  DB: 4 Gaps: 31
US-09-890-475-1 (1-609) x US-09-479-453-29 (1-17612)
QY 14 ThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuPro 33
Db 9754 ACCTCGATCTGCGCGTGCACCAATCGTAGCTACCCCAAGCAACGCGCGGTG 9813
QY 34 LysileValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnPro 53
Db 9814 AAGATCGCGAGATGGAGTGCACCTCGCGTTCATCGCC----- 9852
QY 54 GlnPheLeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLys 73
Db 9853 -----TCGCGCTTCGCGCAACACGCGGTGTCTTCAAG 9885
QY 74 ArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGluAlaIleAspSerLys 93
Db 9886 CAC-----CAGCGTCACGAACACAAATAAGGGTCATCCCTTGCTGAACAG 9930
QY 94 LeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsn-----PheHisGln 110
Db 9931 CCCCGCGCGCGTGGCGGGCTTTTGTGTCACGCGCTTACGTCATCACACTTCTGCGCCAG 9990
QY 111 ProMetLeuSerProProArgAsnAsnValSerValGluThrThrValThrValSerGln 130
Db 9991 GCTGTGTCGCGCT-----GCAAAATCGGCACATGCAGTTTTCGCGC 10032
QY 131 ProSerGlnGluIleValProGluThrSerAsnLysProGluGlyArgMetCysGlu 150
Db 10033 AAATCCGTAACTTGGCGCGCTCGCGCCATGCCATAAAACAAACAAACAGCAAGATG 10092
QY 151 LeuMetCysSerLysGlyLeuArgLysTyrlle-----TyrAlaAsnIleSerAspG 168
Db 10093 GATCTTCTGTTCGGGGACGATCCGCCCATGTCCACCATACCCAGCGCGCTGACGG 10152
QY 168 nAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys-P 188
Db 10153 CTCGGCAAGCCCGCTTGGCGCGCTGCGCTTCGCTTCGCGCAACGCGCGGTG 10212
QY 188 heValLeuAspCysIleGlyLysPheTyrlleGlnGlyArgArgAlaPheThrLysGluS 208
Db 10213 TGCTGCGGAGCCCTTCGCGCAGGTCCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 10259
QY 208 erProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetP 228
Db 10260 -----CCAGCTTGGCGCGCGTGCAGGAG-----G 10284
QY 228 roAspArgGlyLysGlyLysValLysIleGluSerThrIleLysAspGluAla-----G 246
Db 10285 CCCAGCGCTTGGCGCGCGGTGCTGCGCTGCATCTGCTGAGAGCGCGCTTCGAGC 10344
QY 246 luThrAlaAlaValAlaTrpArgLys-----ArgLeuMetThrGluGlyG 261
Db 10345 AGGAGCTGGCGCTTGGCGCGCGGTGCTGCGCTGCATCTGCTGAGAGCGCGAGTGGCGAGGCG- 10403
QY 261 lyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaLysPheG 281
Db 10404 -----ATGGGTGCCAACTTCACTAGCAGCTGGCGCACTCACTCCC----- 10448
QY 281 lyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuLeuLeuLeuLeuLeuLeu 301
Db 10449 -----GAATCCGCGACCTGCTGGAG----- 10469
QY 301 snGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleV 321
Db 10470 -----CAGGAAGATGACGCGCGATCATCGCTGATCAACGCCATCC 10512
QY 321 alGluSerSerIleLysArgGly-----MetHisIleGluAlaLeuGlu-----M 336
Db 10513 TCAGCGAGCGATCAAGCGCGCTCCGACATCCACCTGAAACCTTCGAGAAACGCC 10572
QY 336 etValTyThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheL 356
Db 10573 TGGTGTGCTTTCGCTTCGCGCG-----ATCCTCCGCGAAGTGA 10614
QY 356 euLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaP 376
Db 10615 TCGAACCGCGCGCGAG-----CTGGCG- 10637
QY 376 heLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluT 396
Db 10638 -----GGCTGTGCTGCTCGCGGTCAAGGTCAATGCGG- 10670
QY 396 hrHisLysLeuAspProAla-----LysGluLeuPro----- 406
Db 10671 -----CGCTGACATCCCGAAGCGCGTACCGCAGGCGCGGTATTCGCTCAAGG 10725
QY 407 -----GlyTrpGlnIleLysGluGlnIleValSerLeu-----GluLysA 420
Db 10726 TCGCGGTGCGAGGTGGATATCCGCGTCTCCACCTGCGTTCGCGCAACGCGAGCGGG 10785
QY 420 spThrIleGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeu----- 437
Db 10786 TGGTCTGCTGCTG-----CTCGAAGCGCGCGCGCGCTGCTGCTCAGCATC 10836
QY 438 -----MetGluGluAlaAlaLeuAlaLysArgMetTyrlleGlnGlnIleLysArgPro- 455
Db 10837 TGGCGATGAGCGAGCGC-----GACCGCGCTGCTCGACACCACTGCGCAAGCGCG 10890
QY 456 -ArgLeuSerProMetGluMetProProValThrSerSerSerTyrlleTyrlleAr 475
Db 10891 ACGGATCATCTAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10950
QY 475 GAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSe 495
Db 10951 GCCT--GGTCACTTCAACGA----- 10969
QY 495 rSerTyrlleGlyProSerThrSerPheProHisArgSerArgArgSerProGluTyrlMe 515
Db 10970 -----CGCTCGCGCAATATCTCAGGTGGAAGA-----CCGATCGAGT 11010
QY 515 tValPro-----LeuProHisGlyGlyLeuGlyArgSerVa 527
Db 11011 ACTACTGGAAGCATTCGCGCAGACCCAGGTCAACCGCGGTGACATGACCTTCGCGCC 11070
QY 527 lTyrlAlaTyrlGluHisLeuAlaPro-----AsnSerTyrlSerProG 541
Db 11071 GCGGCTGCGCGCATCTTCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11130
QY 541 yHisGlyHisArgLeuHisArgGlnTyrlSer-----ProSerLeuVal-- 555
Db 11131 ACCAGGAGACCGCGCATCTCGCGTCCGCGGTGCGCGCGCGCGCGCGCGCGCGCG 11190
QY 556 -----HisGlyGlnArgHisProLeuGlnTyrlSerProProIleHisGlyGlnGln 573
Db 11191 CCTGCAACCAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11250
QY 573 nLeuProTyrlleGlnArgValTyrlArgHisSerProSer 587
Db 11251 CCTTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11293
RESULT 9
US-08-977-171-1
; Sequence 1, Application US/08977171
; Patent No. 6232112
; GENERAL INFORMATION:
```

APPLICANT: CATCHESIDE, DAVID E.  
TITLE OF INVENTION: REAGENTS AND METHODS FOR DIVERSIFICATION  
NUMBER OF INVENTIONS: OF DNA  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 3100 No. 6232112 West Center, 90 South 7th Street  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/977,171  
FILING DATE: 24-NOV-1997

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Skoog, Mark T

REGISTRATION NUMBER: 40,178  
REFERENCE/DOCKET NUMBER: 10552.13US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
TELEX:

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9775 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA  
US-08-977-171-1

Alignment Scores:  
Pred. No.: 0.0479 Length: 9775  
Score: 131.00 Matches: 128

Percent Similarity: 33.73% Conservative: 69  
Best Local Similarity: 21.92% Mismatches: 189  
Query Match: 4.22% Indels: 203  
DB: 3 Gaps: 23

US-09-890-475-1 (1-609) x US-08-977-171-1 (1-9775)

QY 174 GluileProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysile 193  
DB 1046 GAGCGGCTTTGCTCGGCTTCGATCA-----GACTGCCTC 1084  
QY 194 GlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAla 213  
DB 1085 CG-AGGCTGCACAGTTCTGGAGGAGGCGCAGACAAGAAATATACCCCTTCTTCATC- 1142  
QY 214 ArgGlnValSerLeuLeuLeuLeuSerPheLeu----- 225  
DB 1143 AAGCCGTTCTGGGCGCATCTCGACAGTTTCACAGTTCGCGCGCAAGGCTAACGCC 1202  
QY 226 -----LeuMetProAspArgGly-----LysGlyLysValLys 236  
DB 1203 ATCCCATCTCCCATCCACTGGCTTGACAACAAGAGAGGAGGCGCGGAAGCTTGCC 1262  
QY 237 IleGluSer-----Trp----- 240  
DB 1263 ATCTCCACCATCTCTCGAGGCTCTGGAGTGTGACCGTCCGATGCTCTCTCCACC 1322  
QY 241 -----IleLysAspGluAlaGluThrAla-----AlaValAlaTrp 252  
DB 1323 GTTGTGTTGATGAGCAGCAGACACTCTCTGGGTCTGTGGTCTACAGCAGTGGCGAGAGTGTG 1382

QY 253 ArgLysArgLeuMetThrGluGlyLeuAlaAlaLysMetAspAlaArgGly 272  
DB 1383 AACGAGGCTCTCAGGACACAGACTGGTGTCTATCAGAGCGGAAG-----CGCGGT 1433  
QY 273 LeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeu 292  
DB 1434 CTCTGTACAAGGGTGTCTACTTCCGGA-----GACACTCAG 1469  
QY 293 AspLeuIleArgMetSerGlySer----- 300  
DB 1470 GAGCTGTCGCACTCTCGTTGACTCGGATAACGATGCTCTCAAGTTTGTCTGTAAGCAG 1529  
QY 301 -----AsnGluIleAlaGlyAlaLeu 307  
DB 1530 AAGGGTCTTTCTGCCACATCGATCCGCTGCTTTGGTCAAGGCGCTTCCC 1589  
QY 308 LysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSer----- 324  
DB 1590 AAGCTCGAGCAGACTTTGATTTCGAGGAACAGTCTGCCCGCGAGGGCTCTTACACTGCC 1649  
QY 325 -----IleLysArgGlyMetHisIleGluAlaLeuMetVal 337  
DB 1650 CGTCTCTTCTCCGATGAGAAGCTAGTCCGCGGCAAGATCATGAGGAGGCTCAGGAGCTC 1709  
QY 338 TyrThr-----PheGlyMetGluAspLysPheSerAlaAla 349  
DB 1710 TGACCGCTCAGACCCCGAGAAATCGCTTTGAGGCTGCGATCTCTTCTACTTTGCT 1769  
QY 350 LeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAla----- 366  
DB 1770 CTTACCAGGCGCTTGTCTCGCGGCTTACTCTTTCGATATCGAAAGAGGCTTGACGCC 1829  
QY 366 ----- 366  
DB 1830 AAGAGTGAAGTCAAGCGCAGGAGTGGAGATGCTAAGGTAAGTGGCTGAGAGGAG 1889  
QY 367 -----LysArgLysAlaGlnSerProLeuAlaPheLysGluAlaThrLysGlnLeu 384  
DB 1890 GGCAATCAAGCTGCGGCGTCCGCTCCGCTGCCACTTCGCGCCCTGTCCACCAAGGAGGCC 1949  
QY 385 AlaValLeuSerSerValMetGlnCys-----MetGlu 395  
DB 1950 GCCCAGGAGCACCACCCCTGAGAAGATCACCATGAGAGTTCGACGCCCTCCAAAGGTCTCT 2009  
QY 396 ThrHisLysLeuAspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnLeuVal 415  
DB 2010 ACCGAGGAGCTCGATGCTCTCAAGCTCTCGCAAAAGTCTGCCCATGCCATCTAC 2069  
QY 416 SerLeu-----GluLysAspThrLeuGln 423  
DB 2070 AAGATCATTTGTCCTCCCATCATCGAGGAGCTCCGCAAGACGCGGCAAGGCTGTTCTGTGG 2129  
QY 424 LeuAspLysGluMetGluLysAlaArgSerLeuSerLeuMetGluAlaAlaLeu 443  
DB 2130 TACACTCAGAGTTCGAGAAGGACTACTCTTACTAGCCC-CGTCCTGAGGCGCCCTT 2188  
QY 444 AlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetPro 463  
DB 2189 CCCCAAGAGAGCTTATGC-AGTCCCTGAGGAGACCA---TTGCTGCCATCGAGCTGCT 2244  
QY 464 ProValThrSerSerTyrSerProLleTyrArgAspArgSerPheProSerGln-Arg 483  
DB 2245 TCAGAACATCCCAAGTTCACGCGCCCAAGAGGAGGAGAAC---CCCTCCAGGTGG 2301  
QY 483 GAspAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrse 503  
DB 2302 AGACCATGCCCGGTGTGTCTGCGCGGTTTCTCTGTCCTCATCGAGGCGCGGTGCT 2361  
QY 503 rPheProHisArgSerArgSerProGluTyrMetValProLeuProHisGlyGlye 523  
DB 2362 ACATCCCGCGGTGTACCGCGTCTCTCCCGAGCAC-----TGCCCTTATGCTGGGTG 2412

QY 523 uGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisG1 543  
Db 2413 TTCCCGC-----CATGG 2424  
QY 543 yHisArgLeuHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuG1 563  
Db 2425 TCGCGCGCTG-----CAACAGATTGTGTTCG 2451  
QY 563 nTyrSerProPro-----lleHisGlyGlnGlnLeuPro----- 575  
Db 2452 CCTCTCCCGCGCGCGAGGAACCATCACTCCCGAGATTGTCCAGTCGCTCAACAAG 2511  
QY 576 -TyrGly---lleGlnArgValTyrArgHisSerProSer----- 587  
Db 2512 TTGGGGCGAGTCATCGTCTCCCGCGGTGCCAGGGCGTAGTCGCATGCGCTACG 2571  
QY 588 ---GluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerAsnSerLe 606  
Db 2572 GCACCGAGAGCATCACCAAGTGCACAAGATTCTCGGCC-CCGTAACCAAGTTCTCACT 2630  
QY 606 uAspProLys 609  
Db 2631 GCTGCCAAGA 2640

## RESULT 10

US-09-620-312D-372  
; Sequence 372, Application US/09620312D  
; Patent No. 6569662

## GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yundong  
; APPLICANT: Wang, Dunwei  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pf-fl\_genes Version 1.0  
; SEQ ID NO 372  
; LENGTH: 7393  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (119)..(3012)  
US-09-620-312D-372

## Alignment Scores:

Pred No.:	0.0337	Length:	7393
Score:	130.50	Matches:	137
Percent Similarity:	30.80%	Conservative:	86
Best Local Similarity:	18.92%	Mismatches:	274
Query Match:	4.20%	Indels:	227
DB:	4	Gaps:	29

US-09-890-475-1 (1-609) x US-09-620-312D-372 (1-7393)  
QY 1 MetSerAsnTyrPro-----ProThrValalaAlaGlnPro 12  
Db 859 ATGCCCAACAGTCCCAACTCAAGCCCATTCAGCCCAAGCCCACTGTTATGGAGAACCT 918  
QY 13 ThrThrThrAlaAsnProLeuLeu-----GlnArgHisGlnSerGluGlnArg 28  
Db 919 ---TTACACAGTCAACCCCTGCTTGTACTCCAGCCAGGCAAGAAAAGAAAGACAAAAA 975  
QY 29 ArgArgGluLeuProLyslleValGluThrGluSerThrSerMetAspIleThrIleGly 48  
Db 976 AGAAGGAAATCTTCAAGGAACCTTGAAGTCTCTGACCCCTGGGAAGGTGTGTCGACA 1035  
QY 49 GlnSerLysGlnProGlnPheLeuLysSerlleAspGluLeuAlaAlaPheSerValala 68  
Db 1036 GAGGAAGGCAAAAGGCCATTC----- 1056  
QY 69 ValGluThrPheLysArgGlnPheAspAspLeuGlnLysHisIleGluSerlleGluAsn 88  
Db 1057 -----AGGAATCTTCAGGAATGGATGAAATGAGGGGCTCTTAAT 1101  
QY 89 AlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnPhe 108  
Db 1102 GGCTCATCAGACCCCAAGCCGACTGGCTAGCATCAAGGCTGAAGCCGACAAGATC 1161  
QY 109 HisGlnProMetLeuSerProArgAsnAsnValSerValGluThrThrValThrVal 128  
Db 1162 TACAGTTTCAGCAATGCCCCAGCCCTTCCATTGGAGGAGTAGCGCCCTGAAAC 1221  
QY 129 SerGlnProSerGlnGluIleValPro-----GluThrSerAsnLysProGluGly 145  
Db 1222 ACTACCCCTACTCAGCCCTGACTCCCTTACATGTGTGACCCAGACATGAGCTGAAGCC 1281  
QY 146 GlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIle 165  
Db 1282 AGC-----TCAGTCAAAACCAACAGCCCTGCTACTCTGACATC 1320  
QY 166 SerAspGlnAlaLys-----LeuMetGluGluIleProSer-----AlaLeu 179  
Db 1321 TCTGATGTGGGAGGATGGGAGGGCAAGGTAGACAGTGTCAATCAAGGACGCCGNA 1380  
QY 180 LysLeuAlaLysGluProAlaLysPheValLeu-----AspCysIleGly 194  
Db 1381 CAGTTGGTTAAAGAAGGGGCTAAGAAAACTCTTTTCCCTCCCTCAGCCTCAGACAAAGAC 1440  
QY 195 LysPheTyrLeuGlnGlyArgArgAlaPheThrLysGluSer----- 208  
Db 1441 TCACCATATTACCAAGGCTTTGAGAGATTACTATTCTCCAAGTTATGCACAGTCCAGCCCT 1500  
QY 209 -----ProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPhe 224  
Db 1501 GGGGCTCTGAACCCCAAGCCAGCCAGCCAGCAGCA-----GTGAGAGCCAG 1542  
QY 225 LeuLeuMetProAspArgGlyLysGlyLysValLyslleGluSerTrpIleLysAspGlu 244  
Db 1543 GCCCTGAAGACAAAAAGGATGAGAACCTCAGACATAGAGGGAAGTGAAGAACCAT 1602  
QY 245 AlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGlyGlyLeuAla 264  
Db 1602 ----- 1602  
QY 265 AlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSer 284  
Db 1603 -----ATCTGTGAAGAAAAAGAGGCC 1623  
QY 285 AsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIleAla 304  
Db 1624 GAGCTGAGCAGTTCAGCTCAGCAGCCCTCGGTTCATCCAGCAGCGTCCCAATATGTACATG 1683  
QY 305 GlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSer 324  
Db ----- 324



Db 1684 CAGTCCCTGTACTACACAGCATGCTATGTATGCTACCCCTATGGCTACAGCGACAGAGT 1743  
Qy 325 lIeLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp 344  
Db 1744 TACCAC-----ACCAGCTTCTGAGCACTAACACGGCTTACCGGCACGACATAGAGAA 1797  
Qy 345 LysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGlu 364  
Db 1798 -----CAGCAGAAACGCCAGAGCTTAGAG 1821  
Qy 365 ArgAlaLysArg-----LysAlaGlnSerProLeuAlaPheLysGluAlaAla 380  
Db 1822 CAGCAGCAGCGGGAGTGACAAAGAGCAGAGATGGCCCTGAAGAGCGGGAGCGCAGCA 1881  
Qy 381 ThrLysGln-----LeuAlaValLeuSerSer 389  
Db 1882 CTCAGGAGAGTGGAAGCAGAAAGCGTCATTCACCACTCTCACCAAGGCCCGCCAGC 1941  
Qy 390 ValMetGlnCysMetGluThr----- 396  
Db 1942 CTGACAGACCTGGTGAAATTCAGACCTGGCAAGGCCAAGGAGCCAGGGGCTGACCCAGCC 2001  
Qy 397 -----HisLysLeuAspProAlaLysGluLeuProGlyTyrGlnIleLys 411  
Db 2002 AAATCAGTCATCTCCCAAGTTAGATGACTCTTCAAACTCCCGGC-----CAGGCCCT 2058  
Qy 412 GluGlnIleValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLys 431  
Db 2059 GAAGGCCCTAAAGTGAAGCTAGTATGCCAGCCACCTAAGCAAGGAGCGCTCTGAGGCC 2118  
Qy 432 AlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGln 451  
Db 2119 AAGCAGGTGTGATGTGGTGCAGCAGCAGAGATGGATCAATCTCTGTACGACAG 2178  
Qy 452 lIeLysArgProArgLeuSerProMetGluMetProProValThrSerSerTyrSer 471  
Db 2179 GAGCAGAGCCCGGATGGACATATGTTATCT-----GCCAAGTACTCA 2226  
Qy 472 Profile-----TyrArgAspArgSerPheProSerGln 482  
Db 2227 GACATCAGTCAGAGATGAGCGGTGGAGGAGCGGCGCCGCAAAATTGAAGAGGAA 2286  
Qy 483 ArgAspAspGlnAspGluIle----- 490  
Db 2287 AGGAGTCGGAGTAAGCACTCTCTCCCAAGGAAGATGGAAAGAACACAGTAGTGAC 2346  
Qy 491 -----SerAlaLeuValSerSerTyrLeuGlyProSerThr 502  
Db 2347 TGCAGCTGCCACGCTCAGAGAGTCTCGCCTTGGGAGCAAGGAGCCCGCCCAAGTGC 2406  
Qy 503 SerPhePro-----HisArgSerArgArgSerProGluTyrMet 515  
Db 2407 CATGTGCTGTCTCTCCCACTTACCAGCACCAGTCC-----TAC 2448  
Qy 516 ValProLeuProHisGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaPro 535  
Db 2449 ATCCCTCATACACGCG-----TATTCCTACAGT----- 2478  
Qy 536 AsnSerTyrSerProGlyHis---GlyHisArgLeuHisArgGlnTyrSerProSerLeu 554  
Db 2479 CAGTCTTACACCCCAACACCCCGAGTACCG-----AGCATGCCTGCT 2523  
Qy 555 ValHisGlyGlnArgHisProLeuGlnTyrSerProProIleHisGlyGlnGlnLeu 574  
Db 2524 GTGATGATCAGAACTACCCAGGTTCTTACCTGCCTTCCAGCTACTCTTTTCC----- 2577  
Qy 575 ProTyrGlyIleGln----- 580  
Db 2578 CCATATGGCAGCAGAGTCTCAGTGGTGAAGTGTGACAGGCGCAGGAGCCCGCCAGT 2637  
Qy 581 ValTyrArgHisSerProSerGluAlaArgTyrLeuGlyLeuSerAsnGlnArgSerPro 600  
Db 2638 GTGACTTGTGAATCCAGCTCAGAGTCCAAAGCCCTGGACATCTTGCAGCAGCATGCCAGT 2697

Qy 601 ArgSerAsnSer 604  
Db 2698 CACTACAGAGC 2709

## RESULT 11

US-09-620-312D-1020  
; Sequence 1020, Application US/09620312D  
; Patent No. 6569662

## GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 1020  
; LENGTH: 3324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (168)..(1625)  
US-09-620-312D-1020

## Alignment Scores:

Pred. No.: 0.0126 Length: 3324  
Score: 129.00 Matches: 100  
Percent Similarity: 37.95% Conservative: 81  
Best Local Similarity: 20.96% Mismatches: 194  
Query Match: 4.16% Indels: 102  
DB: 22 Gaps: 22

US-09-890-475-1 (1-609) x US-09-620-312D-1020 (1-3324)

Qy 47 lIeGlyGlnSerLys---GlnProGlnPheLeuLysSerIleAsp-----Glu 61  
Db 243 CTCTCTCAAGAGAGTGTGCAAGAGCAACATCTTCTCACCACAGGGGTTCACCTAGAT 302  
Qy 62 LeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspLeuGlnLys 81  
Db 303 ATAGCTTCAACTCTCTGGATCAAGAAATTTTATTAAAGTCTTAAACTGAAATTGAAGAA 362  
Qy 82 HisIleGluSerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValValLeu 101  
Db 363 GAGCTAAATCTCTGGACAAAGAAATTTCTGAAGCTTCCAGCAGCAGAGCTTT----- 416  
Qy 102 AlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnValSer 121  
Db 417 -----GACCGTCACACTTCTCCAGTGTTCAGCCCTGCCCAATCCAGAAAGCTCA 464

```
QY 122 ValGluThrThrValThr-----ValSerGlnProSer 132
   :|||
   :|||
Db 465 ATGGAAGACTGCTTGGCCCATCTTGGAGAAAAGTGTCCAGGAACCTGAAGAGCCCTCTC 524
   :|||
QY 133 GlnGluIleValProGluThrSerAsnLysPro-----GluGlyGlyArgMetCys 149
   :|||
Db 525 CATTAAGCATTTGCAAACTCTCTGAGCCGCGAGTGCATATCAGGCAATTCGGGAATGT 584
   :|||
QY 150 GluLeu-----MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSer 166
   :|||
Db 585 ACATGGAGACACACATTCATCGCCAGCGCTGGAATAAGATTTGGTGCCTCTGGTTTG 644
   :|||
QY 167 AspGlnAlaLysLeuMetGluIleProSerAlaLeuLysLeuAlaLysGluProAla 186
   :|||
Db 645 CTACGACAAATGCTTTTGAA-----TTGACAGACTTGTTCAGAACCTTTG 692
   :|||
QY 187 LysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgAla----- 203
   :|||
Db 693 AGCGCACTGCTGCAG---TTTGGCGTGACATACCTGGAGGACTATTGGCAGAGTACATC 749
   :|||
QY 204 -----PheThrLysGluSerProMetSerSerAlaArg 214
   :|||
Db 750 ATTCAAGCAAGTGGCTGGGCACTGTGTGTAGTCTTGTAGTCAGAGGAGGAGGAATACCT 809
   :|||
QY 215 GlnValSerLeuLeuLysSerPheLeuLeuMetProAspArgGlyLysGlyLys 234
   :|||
Db 810 GGAATCACTGCAGAGATAGCAATGACATTTACATCTCTGCCCGGCACTCTGGACAA 869
   :|||
QY 235 ValLysIleGluSerTyrIleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLys 254
   :|||
Db 870 GTCAGT-----CCCCAGAGTCTCCAACTGTGACCACTTCTCTGGCAGTCT 914
   :|||
QY 255 ArgLeuMetThrGluGlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeu 274
   :|||
Db 915 GAGAGCTTACCTGTG---TCACTGTGGCTAGCCAGAGTTGGCACAGAAAGCCTGCCA 971
   :|||
QY 275 LeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeu---LeuAsp 293
   :|||
Db 972 GTGTCACTAGGC-----CCTGAGTCTCTGGCAGCAGATTGCAATGGATCCTGAA 1019
   :|||
QY 294 LeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeu 313
   :|||
Db 1020 GAAGTGAAGAGCTTAGACAGCAAC-----GGAGCTGGAGAGAGAGTGAG--- 1064
   :|||
QY 314 ValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAla 333
   :|||
Db 1065 -----AACAACTCTCTTAATCTGACATTTGTCAGCTGGAGAA 1103
   :|||
QY 334 LeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThr 353
   :|||
Db 1104 GAAGAGGTG-----CCCGAGGGCATGGAAGAGCTGCTGGGCTTCTGTGCTCTGCCA 1157
   :|||
QY 354 SerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerPro 373
   :|||
Db 1158 CGG-----CGGAGCTGCAAGNGGCATCTCTGAAGCCCGAGCTCCC 1199
   :|||
QY 374 LeuAlaPheLysGluAlaAlaThrLysGlnLeuValLeuSerSerValMetGlnCys 393
   :|||
Db 1200 TTGCTTCCACATATCACTGCCACCTCCCTGCTGGGGGACAGGGAACTGTGACACAGAAGTG 1259
   :|||
QY 394 MetGluThrHisLysLeuAspProAlaLys-----GluLeuProGlyTyrGln 409
   :|||
Db 1260 ATCAGATTGGAATATCCAGCCCTGCTACATCTCTGTTGTAGAACCTTGATGAAGAGAG 1319
   :|||
QY 410 IleLys-----GluGlnIleValSerLeuGluLysAsp 420
   :|||
Db 1320 GTGAAAGCAGCAACAACCTGAACCTACTGAAGTGGAGGAGGTGTCCCGCACTGGAACCC 1379
   :|||
QY 421 ThrLeuGlnLeuAspLysGluMetGluIleLysAlaArgSerLeuLeuMetGluGlu 440
   :|||
Db 1380 ACAGAAACGCTGCTGAGTGAGAAGAGAGATAACCCAGGGAGAGAGCCCTTGTGGAGAG 1439
   :|||
QY 441 AlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgPro----- 455
```

```
Db 1440 CTGTCCCTCGCCAGCGAG-----AAGAAGCCGTGCCCGCTGTGAG 1481
   :|||
QY 456 -----ArgLeuSerProMet---GluMetProValThrSerSer 468
   :|||
Db 1482 GGCAAGTCTAGACTGTCCCCCGCGGTGAGATGAAGCCCATCGCTGTCT 1532
   :|||
RESULT 12
US-09-098-901-1
; Sequence 1, Application US/09098901B
; Patent No. 6218144
; GENERAL INFORMATION:
; APPLICANT: Scott, Matthew
; APPLICANT: Sleson, John C.
; TITLE OF INVENTION: Costal2 Genes and their Uses
; FILE REFERENCE: SUN-65P
; CURRENT APPLICATION NUMBER: US/09/098,901B
; CURRENT FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: 60/051,347
; EARLIER FILING DATE: 1997-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3605
; TYPE: DNA
; ORGANISM: D. Melanogaster
US-09-098-901-1

Alignment Scores:
Pred. No.: 0.0205 Length: 3605
Score: 127.50 Matches: 150
Percent Similarity: 35.48% Conservative: 103
Best Local Similarity: 21.04% Mismatches: 244
Query Match: 4.11% Indels: 218
DB: 3 Gaps: 34

US-09-890-475-1 (1-609) x US-09-098-901-1 (1-3605)
QY 18 ProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIleValGlu 37
   :|||
Db 1492 CCATTTCTGGAGCCGAGGAGCCCGAGGAGTGAATCTTGAAGCAGCCAAATTCGGAGTCG 1551
   :|||
QY 38 ThrGluSerThrSerMetAspIleThrIleGlnSerLysGlnProGlnPheLeuLys 57
   :|||
Db 1552 CCAAACTCCGACAAACGAAACGACACAGCAATAGTCGATCGCCCGCATCTGGACGAC 1611
   :|||
QY 58 SerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAsp 77
   :|||
Db 1612 AGCATAGAAGTCTA-----ATGGAAGAGTTTCCGACACAAACAGAC 1653
   :|||
QY 78 Asp-----LeuGlnLysHisIleGluSerIleGluAsn 88
   :|||
Db 1654 GCTCTTATATTGAAAACACACGCTGAATATCTATCCAGCATCCGAAGCGGTATGCAA 1713
   :|||
QY 89 AlaIleAspSerLysLeuGluSer-----AsnGly-----Val 99
   :|||
Db 1714 AGCCAAGACCGGAGATTGAGGCGACAGCCGACAGAAAGATGGTGTATGATCGAAAGTC 1773
   :|||
QY 100 ValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsn 119
   :|||
Db 1774 AGCATTTGGCAGTCGACAGGAGAAAGTGT---CAGCCAGGTGTAGCTTAAGTACTGTCTGAG 1830
   :|||
QY 120 ValSerValGluThrThrValThrValSerGlnProSerGlnGluIleValProGluThr 139
   :|||
Db 1831 CTTGCCATCTTAATCGGCTAGCTTCCACAGACGCGCCTCCGCCCATCATCTGAGTCG 1890
   :|||
QY 140 -----SerAsnLysProGluGlyGlyArgMetCysGluLeuMetCys 153
   :|||
Db 1891 GTCGTCGATCCTCTGAAAGTCTTTCGGCGGAGAGNATCCGTCAGCGGCTCTCGCTGCC 1950
   :|||
QY 154 Ser-----LysGlyLeuArgLysTyrIleTyrAla 163
   :|||
Db 1951 GCGCGCCCACTGCTCTCTATTGAACAGCTGCAGAAAAAATTCGCAAACTGTGTC---GCT 2007
```

QY 164 AsnIleSerAspGlnAlaLysLeuMetGluGluIle-----175  
 Db GAGATCGAGGCAACAAACGACGACGTTACGGGAATCGAAGAACAAATCCAGGTAACAA 2067  
 QY 176 -----ProSerAlaLeuLys-----Leu 181  
 Db ATATAATCGCGCAATG-GTCAAGAACAGCGATACACGACGCCATCGCAAGCAAGATTT 2126  
 QY 182 AlaLysGluProAlaLysPheValLeuAspCys-----IleGlyLys 195  
 Db CACAGAAACGTCGCAAACTTGAGCGCGAGTGCACAGGCCAAGACGATGAGTTAG 2186  
 QY 196 PheTyrLeuGlnGlyArg-----201  
 Db CGCGTAGTTCAAGCGCGGGTCAGTCGAGATTGAGCGATGGACCACGATAATCGGACAT 2246  
 QY 202 -----ArgAlaPheThrLysGluSerProMet-----SerSerAlaArg 214  
 Db CTCGAGCGTCGACTAGAAACCTCAGCTCAATGAAGCATTATGCGGGTGAGCGGACAG 2306  
 QY 215 GlnValSerLeuLeu-----IleLeuGluSerPheLeuLeuMetProAspArgGly 231  
 Db AAGGTGAAGACGACTACAGCAATCGGTGGCGAGTCGGGAAACAGCGCGATGATTACAG 2366  
 QY 232 LysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAla 251  
 Db AAA-----AAGCTCGAAGAGTGCAAGCTGCGCTGCGAGATGGAGCGGAGCTGTC 2423  
 QY 252 TrpArgLysArgLeuMetThrGluGlyLysLeuAlaAlaGlu-----LysMet 268  
 Db CTACGAGATCCAGGAGACTGGCAAGAGCTAGTGAAGCGCAAGTTCTCCGAGCAA 2483  
 QY 269 AspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSer 288  
 Db CAAGGCCCGCGTAAAGCGAGTACAGCT-----AGGATT 2519  
 QY 289 ThrAspLeuLeuAspLeuIleArgMetSerGlySerAsn-----GluIleAlaGlyAlaLeu 307  
 Db AGCGACCTTATACATTTAGCGAGAGTCGGATACCTGGAGAGCAGCGCGGACCA 2579  
 QY 308 LysArgSerGlnPheLeu-----ValProMetValSerGlyIleValGluSerSer 324  
 Db GAAACAGCAGGACCTTCGTCATGATCGCACTTTCGCGGAACCTCGTGACTTGTG 2639  
 QY 325 IleLysArgGlyMetHis-----330  
 Db TTGAAGAACGCTGTCATTTGGACCGCAAACTTAAGCGGACAGGTCTGACGCCAAAG 2699  
 QY 331 -----IleGluAlaLeuGluMetValTyrThr 339  
 Db GAGGAGCGCAAGCTGCTCGAGTGGATGAGGCGCATCGAGGCCATAGATGCGGCCATAGAA 2759  
 QY 340 PheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSer 359  
 Db TTCAAGAACGATGATACACGCGGCCCGCTCCATCGACACGAGGAC---CGAATTGAG 2816  
 QY 360 LysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAla 379  
 Db CGGAGAGGGAGAACAGATGCTGATGGCAGCGCTAAATCGTCTCTCAACGAGGAGATG 2876  
 QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCys-----393  
 Db CGAACACTTCTGTACAAATFACTTCAGAAAGTTATCGATTTCGCGGACTTTCAGAAAG 2936  
 QY 394 MetGluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGluIleLysGluGln 413  
 Db CTGAGCTCAGCTGCTGTCAGTTGGCGTGGAGCGGATGCTCGGAGTGGAGGAGCGT 2996  
 QY 414 IleValSerLeu-----GluLysAspThrLeuGlnLeu 424  
 Db GTTCTGTCCAATGCGGTGCGCCAGGCTAGACTGGAAGCGAAGCAAGCAAGTGGTGTGCTG 3056

QY 425 AspLysGluMetGluGluLysAlaArgSerLeuSerLeuMet-----Glu 439  
 Db CAGCGCCGACGACGAATGAAA-----CTCATTGTGATGCTGCGTCACATGCGGAG 3107  
 QY 440 GluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerPro 459  
 Db GAAAGCTCGGCCAGTTCGGCCAGCTACGAGAACGAGCTTTGGCCCT-----GCC 3158  
 QY 460 MetGluMetProProValThrSerSerSer-----TyrSerProIleTyrArgAsp 476  
 Db TGTGTGCGCGCGCGCTGAGCTAGTATTTCGACTACGATCATTTCTACAAAGGT 3218  
 QY 477 ArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSer 496  
 Db GCGCGCATCCAGCAAG-----GCACTGATC-----3245  
 QY 497 TyrLeuGlyProSerThrSerPheProHisArgSerArgSerProGluTyrMetVal 516  
 Db -----AAAGCCCAAG-----3257  
 QY 517 ProLeuProHisGlyGlyLeu-----GlyArg 525  
 Db CGATGCCCGCTCGCGCTAGCAAAATACAGGACAAAGCAACGACGCGGAGCG 3317  
 QY 526 SerValTyrAlaTyrGluHisLeu-----AlaProAsnSerTyrSer 539  
 Db AACATCTTTGCCAAGTTCATGCTCACCAGATATGCTGCGCGCAGCGCGTTC 3377  
 QY 540 ProGly-HisGlyHisArgLeuHisArgGlnTyrSerProSerLeuVal-----HisG 557  
 Db TCAGGTCCACCGCGCGAGGATCCAC-----GGCCTGATTGAGTCAACCCACG 3428  
 QY 557 LysGlnArgHisProLeuGlnTyrSerProProIle 568  
 Db GCCA-CGGCAACCACTACGTCGACCAACCACTG 3462

RESULT 13

US-09-919-172-97  
 ; Sequence 97, Application US/09919172  
 ; Patent No. 6673545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Paris, Mary  
 ; APPLICANT: Turner, Christopher M.  
 ; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
 ; FILE REFERENCE: PA-0036 US  
 ; CURRENT APPLICATION NUMBER: US/09/919,172  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: 60/222,469  
 ; PRIOR FILING DATE: 2000-07-28  
 ; NUMBER OF SEQ ID NOS: 102  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 97  
 ; LENGTH: 10432  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. 6673545 2700132CB1  
 US-09-919-172-97

Alignment Scores:  
 Pred. No.: 0.121 Length: 10432  
 Score: 127.50 Matches: 138  
 Percent Similarity: 33.43% Conservative: 91  
 Best Local Similarity: 20.15% Mismatches: 230  
 Query Match: 4.11% Indels: 227  
 DB: 4 Gaps: 30

US-09-890-475-1 (1-609) x US-09-919-172-97 (1-10432)

QY 11 GlnProThrThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArg---29  
 Db 2717 GAGCCTTCAAAAAACAGTATCCACTGTAAACAGGTGAGGAGTCTACAGATTTCAGGAAT 2776



Qy	220	IleLeuGluSerPheLeuLeuMetProAsp-----	---ArgGlyLys	232
Db	3323	-----AGAAAGTTTAAGAGAGTCTCCAAAGCAGATCTCGACCCAGCAGCCCGTGTAACT		3376
Qy	233	GlyLysValIysIleGluSerTrpIleLysAspGluAlaGluThrAlaala---ValAla	251	
Db	3377	GGAAATGAAGAGTGGCCAAAGACCGCTTAAGGAAGAGGCCAGTCACTAGGAAGACCTGGCT		3436
Qy	252	TrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaAlaGluLysMetAspAlaArg	271	
Db	3437	GGCTTCAAAGAGCTCTTCCAGACACACAGTCCCTCTGGAAGATCAATGACTGATGAGAAA		3496
Qy	272	GlyLeuLeuLeuValAlaLysPheGlyValPro-----	283	
Db	3497	-----ACTACCAAAATAGCTTCAATCTCCACCACAGCAATCAGTGGACACTCCACA	3550	
Qy	284	-----SerAsnPheArgSerThrAspLeu-----	Leu	292
Db	3551	AGCACAAAGCAATGGCTTAAGAGAAGTCTCAGAAAAGCAGATGTAGAGGAAGAAATCTTTA		3610
Qy	293	AspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPhe	312	
Db	3611	GCACCTCAGGAACCTAACACCATCA-----GCAGGG-----	AAAGCCATG	3649
Qy	313	LeuValProMetValSerGlyIleValGluSerSerIleLys-----	ArgGlyMetHis	330
Db	3650	CTTACGCCCAACACGACGAGGATGAGAAAGACATTAAAGCAITTTATGGGAATCCCA	3709	
Qy	331	IleGluAlaLeuGluMetValTyThrPheGlyMetGluAspLysPheSerAlaAlaLeu	350	
Db	3710	GTGCAGAACTGACCTCGCAGGAACCT-----	3736	
Qy	351	ValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAla	370	
Db	3737	-----TTACTGCGCACAAAGACAGACGTACAGACTCTTAAGGAAAAGGCC	3781	
Qy	371	GlnSer-----ProLeuAla---PheLysGlu-----	378	
Db	3782	CAGGCTCTAGAACACCTGGCTGGCTTTAAAGAGCTTCTCCAGACTCTGGTCAACCCGAG	3841	
Qy	378	-----	378	
Db	3842	GAATTAGTGCTGCTGGTAAACCACTAAATACCTCGAGCTCTCCAGTCCAGACCCCA	3901	
Qy	379	-----AlaAlaThrLysGln-----	383	
Db	3902	GTGGACACCCCAACAAGCACAAAGCAACGACCCCAAGAGAAGTATCAGGAAGACAGATGTA	3961	
Qy	384	-----LeuAla-----ValLeuSerSerValMetGlnCysMetGluThr	396	
Db	3962	GAGGGAGAACTCTTAGCGTGAGGAATCTAATGCCATCAGAGGCAAGCCATGCAACG	4021	
Qy	397	-----	His	397
Db	4022	CCTAAACCATCAGTAGTGAAGAGAAAGACATCATATTTTGGGAACCTCCAGTGCAG	4081	
Qy	398	LysLeuAspProAlaLysGluLeuProGlyTyTrpGlnIleLysGluGlnIleValSerLeu	417	
Db	4082	AAACTGGACCTGACAGAACTTAAACCGCAGCAAGAGACGGCCCAAACTCTCAAGGAA	4141	
Qy	418	GluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaAlaArgSerLeuSerLeu	437	
Db	4142	GAGGCCAGGCTCTGGAAGACCTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCAT	4201	
Qy	438	MetGluGluAlaAlaLeuAlaLysArgMetTyAsnGlnGlnIleLysArgProArgLeu	457	
Db	4202	ACTGAAGAAGACAGTGGCTGCTGGC-----	AAAACTACTAAATG	4240
Qy	458	SerProMetGluMetProProValThrSerSerSerTySerProIleTyArgAspArg	477	
Db	4241	CCCTGGGAATCTTCTCCACCAAGAAATCAGACAGACACCCCAACAGACCAAGAGGACGCC	4300	
Qy	478	SerPheProSerGlnArgAspAspAspGlnAspGluIleSerAlaLeuVal-SerSerty	497	

```
Db 4301 AAGACACCTTTGGAGAAAGGACGTACAGAGAGGCTCTCAGCCCTGAAGAAGCTCA-- 4358
Qy 497 rLeuGlyProSerThrSerPheProHisArgSerArgSerProGluTyrMetValPr 517
Db 4359 -----CACACATCAGGGGAAACACACACAGATATA 4393
Qy 517 oLeuProHisGlyLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsnSe 537
Db 4394 -GTACCA--GGAGGTGAGGATAAAGCATCAACGGCTTTAGGGAACTGCAAAACAGAA 4449
Qy 537 rTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHisGl 557
Db 4450 ACTGGACCCA-----GCAGCAAGTGTACTGCTAG 4479
Qy 557 yGlnArgHisPro 561
Db 4480 CAAGAGGCACCCA 4492

RESULT 15
US-08-466-390-3
; Sequence 3, Application US/08466390
; Patent No. 5686562
; GENERAL INFORMATION:
; APPLICANT: TOKUATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,390
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6306 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..6306
; PUBLICATION INFORMATION:
; AUTHORS: COMPTON, DUANE A
; AUTHORS: SZILAK, ILVA
; AUTHORS: CLEVELAND, DON W
; TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
; TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
; TITLE: SEGREGATION OF PROTEINS AT MITOSIS
; JOURNAL: J. Cell Biol.
; VOLUME: 116
; PAGES: 1395-1408
; DATE: 1992
```

```
US-08-466-390-3
Alignment Scores:
Pred. No.: 0.0587 Length: 6306
Score: 127.00 Matches: 161
Percent Similarity: 34.61% Conservatives: 102
Best Local Similarity: 21.18% Mismatches: 275
Query Match: 4.09% Indels: 222
DB: 1 Gaps: 35

US-09-890-475-1 (1-609) x US-08-466-390-3 (1-6306)
Qy 9 AlaAlaGlnProThrThrThrAlaAsnProLeuGlnArgHisGlnSerGlu----- 26
Db 3331 GCTGCTGCAGCAGACAGAGCCCAACAGGCCCAAGCTTGAAGCAGCTGCGGCGCAGAGGTGAGC 3390
Qy 27 -----GlnArgArgGlu-----LeuProLysIleVal 36
Db 3391 AAGCTGGAACAGCAATGCCAGACAGCAGCAGGCTGACAGCTTGGACGCGCTC 3450
Qy 37 GluThrGluSerThrSer-----MetAspIleThrIleGlyGln 49
Db 3451 GAGGCTGAGCGGGCTCTCCGGGCTGAGCGGACAGTGTCTTGGAGACTCTCGAGGGCCAG 3510
Qy 50 -----SerLysGlnProGlnPheLeuLysSerIleAsp 60
Db 3511 TTAGAGGAGAAGGCCAGGAGCTAGGGCAGCTCAGAGTGCCTTAGCTCGGCCCAACGG 3570
Qy 61 GluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspAspLeuGln 80
Db 3571 GAGTTGGCTGCTTCCGCCAACCAAGGTACAGACCCACAGCAGAGCTGAAGATGAGTGGAA 3630
Qy 81 LysHisIle-----GluSerIleGluAsnAlaIleAspSerLysIle 94
Db 3631 GCCCAGGTGGCGCGCGCCGCAAGAGCTGAGAGGAAATAGCTCATCAGCAGCTTG 3690
Qy 95 GluSerAsnGlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMetLeuSer 114
Db 3691 GAGGAGGAGGTGCTCCATCTG-----AATGCGCAGGTCTCTGGAG 3729
Qy 115 ProProArgAsnAsnValSerValGluThrThrValThrVal-----SerGlnProSerGln 133
Db 3730 AAGGAGGGGAGAGCAGCAGAGTGAAGCGGCTGTGATGCGCGAGTCAAGAGCAGCAG 3789
Qy 134 GluIle-----ValProGluThr 139
Db 3790 AAGCTGGAGGAGAGCTGCGCTGTCAGGCGAGACAGACAGCAGTCCAGAGCTG 3849
Qy 140 SerAsnLysPro---GluGlyGlyArgMetCys----- 149
Db 3850 CAGAACGCGAGCTCTGCTCTGCGGGAGGAGGTGCGAGCGCTCCGGAGGAGGCTGAGAAA 3909
Qy 150 GluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAla 169
Db 3910 CAGCGGTGGCTTCAGAGAACTTGGCGAG-----GAGCTGACCTCAGAGCT 3957
Qy 170 LysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGlu----- 184
Db 3958 GAGCGTGGCGAGAGCTGGCGCAAGATTGAAGGCGTGGCAGGAGAGTCTTCCAGAAA 4017
Qy 184 ----- 184
Db 4018 GAGCAGGCCCTCTCCACCTCGAGCTCGACACACACAGCAGCAGGCGCTGGTGGAGT 4077
Qy 185 -----ProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg 202
Db 4078 CTGTGTCAGCTTAAGCACCTCTGCGCAG-----CAGCTGACGCGCAGCAG 4122
Qy 203 AlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIle----- 220
Db 4123 GCCGCTGCCGAGAAACGCCACCGTCGTGAGGAGTGGAGCAGACAGCAAGCGCGCTGGGGGA 4182
Qy 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLys---ValIysIleGluSer 239
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 28, 2004, 04:51:41; Search time 523 Seconds

(without alignments)

4200.664 Million cell updates/sec

Title: US-09-890-475-1

Perfect score: 3104

Sequence: 1 MSNPPVTAQAQPTTANPLL.....RYLGSLNQRSPRSNSLDPK 609

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delpop 6.0, Delpext 7.0

Searched: 253733 seqs, 180373377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEVEL=xl

-Q=/cgm2\_1/USPTO\_spool\_p/US09890475/runat\_24022004\_135314\_5422/app\_query.fasta\_1.775

-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blowum62

-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=2000000000 -USER=US09890475@cgm2\_1.107@runat\_24022004\_135314\_5422

-NCFU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA:

1: /cgm2\_6/ptodata/2/pubnpa/US07\_PUBCOMB.seq:

2: /cgm2\_6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq:

3: /cgm2\_6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq:

4: /cgm2\_6/ptodata/2/pubnpa/US06\_PUBCOMB.seq:

5: /cgm2\_6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq:

6: /cgm2\_6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq:

7: /cgm2\_6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq:

8: /cgm2\_6/ptodata/2/pubnpa/US08\_PUBCOMB.seq:

9: /cgm2\_6/ptodata/2/pubnpa/US09A\_PUBCOMB.seq:

10: /cgm2\_6/ptodata/2/pubnpa/US09B\_PUBCOMB.seq:

11: /cgm2\_6/ptodata/2/pubnpa/US09C\_PUBCOMB.seq:

12: /cgm2\_6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq:

13: /cgm2\_6/ptodata/2/pubnpa/US10A\_PUBCOMB.seq:

14: /cgm2\_6/ptodata/2/pubnpa/US10B\_PUBCOMB.seq:

15: /cgm2\_6/ptodata/2/pubnpa/US10C\_PUBCOMB.seq:

16: /cgm2\_6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq:

17: /cgm2\_6/ptodata/2/pubnpa/US60\_NEW\_PUB.seq:

18: /cgm2\_6/ptodata/2/pubnpa/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match Length	DB ID	Description
------------	-------	--------------	-------	-------------

1	773.5	24.9	2368	12	US-10-424-599-59479	Sequence 59479, A
2	2041	12.0	2041	12	US-10-425-114-21596	Sequence 21596, A
3	362.5	11.7	2206	12	US-10-424-599-5616	Sequence 5616, Ap
4	361	11.6	2089	12	US-10-425-114-30101	Sequence 30101, A
5	340	11.0	1931	12	US-10-425-114-35432	Sequence 35432, A
6	325	10.5	2237	12	US-10-425-114-32180	Sequence 32180, A
7	325	10.5	2300	12	US-10-425-114-5736	Sequence 5736, Ap
8	309.5	10.0	1780	12	US-10-424-599-43984	Sequence 43984, A
9	297	9.6	2199	12	US-10-424-599-49586	Sequence 49586, A
10	294	9.5	1812	12	US-10-424-599-95969	Sequence 95969, A
11	292	9.4	2393	12	US-10-424-599-49595	Sequence 49595, A
12	265.5	8.6	1772	12	US-10-425-114-34508	Sequence 34508, A
13	234.5	7.6	1524	9	US-09-938-842A-917	Sequence 917, App
14	234.5	7.6	1524	11	US-09-938-842A-917	Sequence 917, App
15	188.5	6.1	1454	12	US-10-425-114-35315	Sequence 35315, A
16	178	5.7	936	12	US-10-424-599-59477	Sequence 59477, A
17	174	5.6	1420	12	US-10-424-599-16050	Sequence 16050, A
18	172	5.5	1896	12	US-10-425-114-14462	Sequence 14462, A
19	161.5	5.2	1179	12	US-10-425-114-21522	Sequence 21522, A
20	151	4.9	3312	15	US-10-445-735-1	Sequence 1, Appli
21	148.5	4.8	3048	14	US-10-317-835-14	Sequence 14, Appl
22	148.5	4.8	3824	10	US-09-291-417-20	Sequence 20, Appl
23	148.5	4.8	4535	14	US-10-317-835-12	Sequence 12, Appl
24	148.5	4.8	4536	14	US-10-317-835-11	Sequence 11, Appl
25	148.5	4.8	4620	14	US-10-177-293-253	Sequence 253, App
26	148.5	4.8	4620	14	US-10-317-835-10	Sequence 10, Appl
27	145.5	4.7	2760	10	US-09-893-519A-146	Sequence 146, App
28	145	4.7	2187	15	US-10-023-634-17	Sequence 17, Appl
29	145	4.7	3045	14	US-10-240-965-212	Sequence 212, App
30	144.5	4.7	3003	13	US-10-317-835-13	Sequence 13, Appl
31	144	4.6	2684	15	US-10-023-634-13	Sequence 13, Appl
32	144	4.6	2708	14	US-10-256-250-2	Sequence 2, Appli
33	144	4.6	2756	9	US-09-960-253-159	Sequence 159, App
34	144	4.6	2756	14	US-10-256-250-1	Sequence 1, Appli
35	142	4.6	3114	15	US-10-205-647A-3	Sequence 3, Appli
36	141	4.5	2020	15	US-10-023-634-15	Sequence 15, Appl
37	140.5	4.5	2609	14	US-10-256-250-3	Sequence 3, Appli
38	140.5	4.5	3587	15	US-10-104-047-1311	Sequence 1311, Ap
39	140	4.5	3718	9	US-09-917-800A-1471	Sequence 1471, Ap
40	136	4.4	6455	15	US-10-373-801-3	Sequence 3, Appli
41	134	4.3	2826	9	US-09-938-842A-1242	Sequence 1242, Ap
42	134	4.3	2826	11	US-09-938-842A-1242	Sequence 1242, Ap
43	132.5	4.3	3913	14	US-10-233-045-12	Sequence 12, Appl
44	132.5	4.3	5142	14	US-10-097-340-42	Sequence 42, Appl
45	131	4.2	9775	9	US-09-751-962-1	Sequence 1, Appli

#### ALIGNMENTS

##### RESULT 1

US-10-424-599-59479

Sequence 59479, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 59479

LENGTH: 2368

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT\_MRT3847\_24720C.1

US-10-424-599-59479

Alignment Scores:



Pred. No.: 1,43e-74 Length: 2368  
Score: 773.50 Matches: 162  
Percent Similarity: 64.38% Conservativity: 91  
Best Local Similarity: 41.22% Mismatches: 108  
Query Match: 24.92% Indels: 32  
DB: 12 Gaps: 5

US-09-890-475-1 (1-609) x US-10-424-599-59479 (1-2368)

QY 57 LysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPhe 76  
DB 195 AAATCGGTGACAACTCAACAGCCTCGCTTCGATACAGCCTCAATAACAGGTAC 254  
QY 77 AspAspLeuGlnLysHisIleGluSerIleGluAlaIleAspSerLysLeuGlu--- 95  
DB 255 GAAGAAATTCAGAAACACCTCGAATTCATCGAACAGCCATTGACAGGAGCCAGGAG 314

QY 96 -----SerAenGlyValValLeuAla 102  
DB 315 CTCCGAGCAGTACGCTCAATCTACTCAAGAACCCGCGAAACCGCGCTCGTTCAATCA 374

QY 103 AlaArgAsnAenAsnPheHisGlnProMetLeuSerProArgAsnAenValSerVal 122  
DB 375 -----GATTCAATCCGAGGCCA 392

QY 123 GluThrThrValThrValSerGlnProSerGlnGluIleValProGluThrSerAsnLys 142  
DB 393 GAAGAAGCAAGCGGTGAG 452

QY 143 ---ProGluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTrpIle 161  
DB 453 GAAGACGAGCTTATACGCTTTGCAAAACGATGATAGCCGAGCGCTCGTAAATACGTG 512

QY 162 TyrAlaIleSerAspGlnAlaLysLeuMetGluIleProSerAlaLeuLysLeu 181  
DB 513 TTAACGCGCTATCCGAAACCGCGCTTCGGAACAGGTACCCGTTGCGCTGAGGAGC 572

QY 182 AlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg 201  
DB 573 GCGGGAAGCCCTCAAGCTGCTGTTGAATGCAATGGAGGTTTTCCTTCAGGGAGC 632

QY 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221  
DB 633 AAAGCTTACAGAAAGAACTCCCGATGTTCCCGCAAGCAGGAGTTTCGGTCTGTTTG 692

QY 222 GluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrpIle 241  
DB 693 GAGCACTACTTCTCTCGCTGCTGCTGGGAATGAGAG---GACGTGGAGGCTTCATTG 749

QY 242 LysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGly 261  
DB 750 AAGAGGAGGCGGATTCGCGCGGCTTCGCTGAGAGAGAGATGTTTGTGAAGCGGGT 809

QY 262 LeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGly 281  
DB 810 TTGCTGAAGGAGCTGAGGTGATGCCAGGGGTTGATTTCTTCGTCGCGCGCTTCGGG 869

QY 282 ValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsn 301  
DB 870 ATTCACAGTGTTCACAGGATGAGATATATACAACTTGTTGTTGTCAGCAATGGCAGA 929

QY 302 GluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleVal 321  
DB 930 GAATTCCTCGATGCGCTCTCCAGTCTCAGCCCTGCTTAAGAGGGTTTCAGATGTTGCA 989

QY 322 GluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGly 341  
DB 990 GATGGGATGATCAAAAGAGGATGCCCGTTAAAGCTGTGATTTGGCTTATACCTTTGGG 1049

QY 342 MetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGlu 361  
DB 1050 TTTGAAGAGAAATATCTTCTCAGACAGCTCTGATCTTCAATTTCTGAGAGTCTGAAGA 1109

QY 362 SerPheGluArgAlaLysArgLysAlaGlnSer---ProLeuAlaPheLysGluAlaAla 380  
DB 1110 ACTTGGAAAGAACCCAAACAGACGCGACGTGATTTCTTAGTCACTGAGGTAGCACAT 1169

QY 381 ThrLysGlnLeuAlaValLeuSerValMetGlnCysMetGluThrHisLysLeuAsp 400  
DB 1170 GAAATAATTTGGCTGCTTTCATCTGTAGTTAAATGTTTGGAGGTCACAGATTGAC 1229

QY 401 ProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIle-ValSerLeuGluLysAs 420  
DB 1230 CCTGTAAATTTCTGCTGCTGGTGGCACTTAAGATAAGATTACTCACTTGGAGAAAGA 1289

QY 420 pThrLeuGlnLeuAspLysGluMetGluGluLysAla 432  
DB 1290 TATTAATGATGCAATAAAAAAATTTGATGAGAGTCA 1326

RESULT 2  
US-10-425-114-21596  
Sequence 21596, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313) B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 21596  
LENGTH: 2041  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3354-037-B3\_FLI  
US-10-425-114-21596

Alignment Scores:  
Pred. No.: 1,63e-30 Length: 2041  
Score: 373.00 Matches: 157  
Percent Similarity: 40.06% Conservativity: 103  
Best Local Similarity: 24.13% Mismatches: 231  
Query Match: 12.02% Indels: 158  
DB: 12 Gaps: 27

US-09-890-475-1 (1-609) x US-10-425-114-21596 (1-2041)

QY 5 ProProThrValAlaAlaGlnProThrThrThr-----AlaAsnPro 18  
DB 123 CCTCTCTCTTCAGCAGCAGCGCGGCGCACACCTCGATATCTTCCCTCGGACCG 182

QY 19 LeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuPro----- 33  
DB 183 -----GAGCGAGCAGCAGCCTCCAGCGCTCCACCGCTCCCGCTCGTTATTGGGA 236

QY 34 -----LysIleValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLys 51  
DB 237 GGATTCGCATCTCTGACATGGAGTCCGTCGAGCTCTTATGAACTCGACAGCTCCAG 296

QY 52 GlnProGlnPheLeuLysSerIleAspGluLeuAlaAla----- 64  
DB 297 ATACAGCAGCTTCAGGAGGCAATTTGCTGAGCTTGAGAGCCAGAGCAGCAGCGTTTCCATGAC 356

QY 65 -----PheSerValAlaValGluThrPheLysArgGln 75  
DB 357 CTCAGTGAAGCAGCTTGAAGACCACTTTCTGTGTTGAACAGTCCCTCAAGAAAG 416

QY 76 PheAspAspLeuGln----- 80  
DB 76 PheAspAspLeuGln----- 80



Qy	13	ThrThrAlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGlu---	31
Db	405	ACAAGGAAGGCTCAGGAGATACTGGAGACGGCAGCAGCTGTTTATGCCAAGAGCAA	464
Qy	32	-----LeuProIlylValGluThrGluSerThrSerMet---AspIleThrIleGly	48
Db	465	GCCACGTGTGCAGAGGCTTCAGGAAAGGGATGCTGCTGATTCGATCTTGAATGCT	524
Qy	49	GlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPheSerValAla	68
Db	525	CGAGAAAGCAGAGGAGGTC-----ACGATAAGTGATTGGCTATTGCTCCAATGGG	578
Qy	69	ValGlu---ThrPheLysArgGlnPheAspAspLeuGlnLysHisIleGluSerIleGlu	87
Db	579	GGTAAGGGGACATTT-----CATGTGGAG-----GAC	605
Qy	88	AsnAlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsn	107
Db	606	AAACCAGTGGATGCT-----GTGTCCTTGCAGCTAATGGTAAT	644
Qy	108	PheHisGlnProMetLeuSerProProArgAsnAsnValSerValGluThrThrValThr	127
Db	645	GTGGAAGAGGTGTACTTCTCTCGAAATCGAAATGTGAGTTG-----	689
Qy	128	ValSerGlnProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArg	147
Db	690	-----TCATTATCCAGACTTGGTAAAA	710
Qy	148	MetCysGluLeuMetCysSerLysGlyLeuArgLysTyrlleTyrlalaHenIleSerAsp	167
Db	711	CTTTGTAAAGAGATGGATGCTGTGGGCTTCACAAATTCATATCTGACACCGTAGAAG	770
Qy	168	GlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLys	187
Db	771	CTTGCTGCTGTGAAGGAGGAAATACCAATGCAATTAAGACAGCTCTTAACGGCGCTGT	830
Qy	188	PheValLeuAspCysIleGlyLysPheTyrlleZeuGln-----GlyArgArgAlaPheThr	205
Db	831	TTAGTTTTAGATCTCTGGAAGGGGTTTACTGTCACAGAAGTGTCAAATCAGGACGTAAAG	890
Qy	206	LysGluSerProMetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeu	225
Db	891	AAGATGCTAACTTACTGGGTGTTCGCGCAACATGTATCATGTGTGTAATGCTCTATGT	950
Qy	226	---LeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGlu	244
Db	951	GATTTCTGAGCACTCAGGTGTGTTCTTAAGTANTTTCAGNAGNATCAGAGACCGG	1010
Qy	245	AlaGluThrAlaAlaValAlaAlaTrpArgLysArgLeu---MetThrGluGlyGlyLeuAla	263
Db	1011	GCTAAGCGAGTGTGTAAGAAATGAAACCCAGATTTGATGCTCTTGACATCGATGTAGC	1070
Qy	264	AlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaCysPheGlyValPro	283
Db	1071	AATGTGAATTCCTTGGAGGCTCATGCAATTTTGCACACTTCTAGCCAGTTTGTATTGCC	1130
Qy	284	SerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIle	303
Db	1131	TCTGGTTTTAATGAGGAGGAGTTATCTAGGCTGATTCCAATGGTATCTCGACGTCCCAA	1190
Qy	304	AlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSer	323
Db	1191	ACTGCTGATTTATGCTGTGCTTGGGTGTGCAGAAAGATGCTGGTGTCAITGAAGTT	1250
Qy	324	SerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrlThrPheGlyMetGlu	343
Db	1251	TTGCGTGAATAGTGGCGGCAAAATGATGCTGTTPAACTTGGCTTTTGCATTGTATACA	1310
Qy	344	AspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPhe	363
Db	1311	GAACAAATTTCAACCATTCCTTTACTGAAGTCTTACTTGAAGGATGCTAGAAAATTTCT	1370

Qy	364	GlueGAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaThrLysGln	393
Db	1371	TCCTCCTGTCAGAGTGTTAACTCATCTCCACATGCACAGATTGATGTTAATGATCCAGAG	1430
Qy	384	LeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLys	403
Db	1431	CTGATTGCACCTTAAAGCTGTAACTCAAGTCGATTAAGATCATATAAATTGAT	1484
Qy	404	GlueLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGln	423
Db	1485	CAGTATCTCT	1523
Qy	424	LeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeu	443
Db	1524	CTGAGAAAGCCCAAGCTGCAAGAGAGGGAA	1568
Qy	444	AlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetPro	463
Db	1569	-----AAGCCTCAACCCAGAGACCCCGTCCAAATGGTGTG-----	1604
Qy	464	ProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArg	483
Db	1605	-----GGATATGTCCA-----	1616
Qy	484	AspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSer	503
Db	1617	-----CGTGTCACTAACATCTTTTCGGACAAAACCTTCC	1649
Qy	504	PheProHisArgSerArgArgSerProGluTyrMet-ValProLeuProHisGlyGlyLe	523
Db	1650	TATGCTAGAGTTGTCGACAGTATCTCTCAATACGTATGTATGACCGACCCCT	1698
Qy	523	uGlyArgSerValTyrAlaTyr-----GluHisLeuAlaPro-----	535
Db	1699	-----TACATGTACCTGCACCAACTGAGAATCATTTGCCCCCTCTCATGAC	1745
Qy	536	-----AsnSerTyrSerProGlyHisGlyHisArgLeuHisArg-----GlnTyr	550
Db	1746	CACCTGCAACATATAACATCTCTCCGCCATGCGCAACTACTTTTGGAAATGGGTATCAGTA	1805
Qy	550	rSerProSerLeuValHis	556
Db	1806	CCAAAGCCTCATATCTCCAC	1824
RESULT 4			
US-10-425-114-30101			
; Sequence 30101, Application US/10425114			
; Publication No. US2004003488A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated Wi			
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 30101			
; LENGTH: 2089			
; TYPE: DNA			
; ORGANISM: Glycine max			
; FEATURE:			
; OTHER INFORMATION: Clone ID: UC-GMROPIC109C08_FLI			
US-10-425-114-30101			
Alignment Scores:			
Pred. No.:		3.56e-29	2089
Score:		361.00	137
Percent Similarity:		39.39%	97



ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: UC-ZMROB73028P10\_FLI  
 US-10-425-114-35432

Alignment Scores:  
 Pred. No.: 6,51e-27 Length: 1931  
 Score: 340.00 Matches: 150  
 Percent Similarity: 39.54% Conservative: 109  
 Best Local Similarity: 22.90% Mismatches: 214  
 Query Match: 10.95% Indels: 182  
 DB: 12 Gaps: 27

US-09-890-475-1 (1-609) x US-10-425-114-35432 (1-1931)

QY	4	TyrProThrValAlaalaGlnProThrThrThrAlaAsnProLeuLeuGlnArgHis	23
DB	51	TGGCCACCCCTCGATCTTTGACCCGAGGCG	89
QY	24	GlnSerGluGlnArgArgGluLeuProLysLeuVal	36
DB	90	GGGGCCAGCAGCGCGCCGCGCTGCTTTGTTATAGGAGCGTTCCGCCATCTCT	149
QY	37	GluThrGluSerThrSer-----MetAspIleThrIleGlyGlnSerLysGlnProGln	54
DB	150	GAAATGGAGTCTGTCGCGCTCTTATGACATCGACA-----AGCTCCAGATACACAG	203
QY	55	PheLeuLysSerIleAspGluLeuAla	64
DB	204	CTTCAGGAGGATTCGCTGAGCTTGAGAGCAGCAGCGCTTTCTGCATGAATTCAGATGG	263
QY	65	-----PheSerValAlaValAlaGluThrPheLysArgGlnPheAspAsp	78
DB	264	AAGCAGCTTGAGACCATCTTCGTGCTTGAACATCCCTCAAGAAAGTTTGACGAG	323
QY	79	LeuGln	80
DB	324	CTAAAGAGCAGGAGAGAGTTCAGGAGACTGTGCAAAATCAGACGATGCTGGAG	383
QY	81	-----LysHisIleGluSerIleGluAsnAlaIleAspSer	92
DB	384	CAGCGGAGGCGAGCTGTTGGCCAGGAATCTGACCTTTTGGAGAGTTCCAGGAGAAG	443
QY	93	LysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMet	112
DB	444	AGGGAAGCTGCATTAGCTATGATCTTTAGCAAGTCCAGGCTTCTTACCTGTTCTCC	503
QY	113	LeuSerProProArg-----AsnAsnValSerValGluThrThrValThrValSer	129
DB	504	ATCAACCAATGAACAAGGCACTGGATAATCTTGTTGTCAGTGGCCTTAAACCGGCTTCT	563
QY	130	GlnProSerGlnGluIle-----ValProGluThrSerAsnLysPro-----GluGlyGly	146
DB	564	GAGGAGAGTGTGACCTTGCAGTATGATATGCTGCAGTGGAGCTCTGTTCTGAGCTGTT	623
QY	147	ArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrlleTyrlleAlaAsnIleSer	166
DB	624	GCACCTCTGTGAGAAATGAATGTGAACGGGCTTCATAGTTTCATATCCGACACAGGAAG	683
QY	167	AspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAla	186
DB	684	AACTTGGCGCATCCGGGAGGAATTTCTAGTGCATCGAAGAAACATCTCATCTTAT	743
QY	187	LysPheValLeuAspCysIleGlyLysPheTyrlle	200
DB	744	GGCGTGTGTGATTCCTGGAAGACTTTTATCTCGAGATAATCTTGTTATAGATGA	803
QY	201	ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeu	220
DB	804	AAA-----AAGATGGTGACCTTCTGGGTGTAGAGGACATGTTTGTGTTG	851
QY	221	LeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysVal	235

DB	852	ATGGAGTCACTT-----GGACAACTGCACCTCTAATGATATA	887
QY	236	-----LysIleGluSerTrpIleLysAspGluAlaGluThr	247
DB	888	ACTTGTCTTCTTTGGAGAGGCACATGCTTCAACTAATATCATAGAGCGCAAGAGC	947
QY	248	AlaAlaValAlaTrpArgLysArgLeu-----MetThrGluGlyGly	261
DB	948	ATTGCATTTTAAAGTGAAGTCCCAAGTTAGATAATCTTGACATTTGATGCTAGCAATGGGAAC	1007
QY	262	LeuAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGly	281
DB	1008	TGT-----CTTGAAGCTCATGCGTTTCTTCAACTCTGCGCAACCTTTGTT	1052
QY	282	ValProSerAsnPheArgSerThrAspLeuLeuAspLeuIle-----ArgMet	297
DB	1053	ATTTCTGCTGAATTCATGAAGATGATGTGCAAAATGCTTCCATATGTAGTGTGCT	1112
QY	298	SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal	317
DB	1113	CGTCAGACACCTGAACCTTTGTCGATGCTGGTGTGTCGAG-----AAATG	1160
QY	318	SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal	337
DB	1161	CCAGGTGTCTGAAGTCTCTGCTGGAAGTGCAGACCTATTGATGCAATTAATTGGCT	1220
QY	338	TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys	357
DB	1221	TATGTGTTTGAAGTCTTCTGAACATTTGAACAGTACACCTTCTTTAAAGCATATCTAAGG	1280
QY	358	MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLys	377
DB	1281	GATGTTAAGAAA-----ATGTCACAGCCAGGAATGTCAAACTTCTCTGGACACAGAA	1337
QY	378	GluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHis	397
DB	1338	GAGATGAATGAACGTGAGTATCTGCTGTAATCTGTCAATTAAGTGCATTAAGAGCAC	1397
QY	398	LysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeu	417
DB	1398	AAACTTGAG-----CAGCAGTACCTT-----GTGATCCACTT	1430
QY	418	GluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeu	437
DB	1431	CAGAAAGGCTTTTGCAGCTAGAGAAAGCAAGGAGCAGACAGAGGATGCTGCTT	1484
QY	438	MetGluGluAlaAlaLeuAlaLysArgMetTyrlleAsnGlnGlnIleLysArgProArgLeu	457
DB	1485	-----GAAGCAGCA-----AAGCCGAGTCCAGAGGCTCTGTGCC	1520
QY	458	SerProMetGluMetProProValThrSerSerTyrlleTyrlleArgAspArg	477
DB	1521	AATGGATCGGCTTT-----GCAACCCGCTGCTACAGGCTTCTGACAG	1565
QY	478	SerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrlle	497
DB	1566	AGCTTCTACTCA-----	1577
QY	498	LeuGlyProSerThrSerPheProHisArgSerArgArgSerProGluTyrlleMetValPro	517
DB	1578	-----GCAACCCAGAGAGGATCCATCAATCTTGGAGAGG-----	1616
QY	518	LeuProHisGlyGlyLeuGlyArgSerValTyrlleAlaTyrlleGlu-----HisLeuAlaPro	535
DB	1617	-----CAGTTTGTGTGCTGCGGTGAGGTTTCATCTCTCCCATG	1655
QY	536	-----AsnSerTyrlleSerProGlyHisGlyHisArgLeu	546
DB	1656	ATGACCTCAGGCTTACAGATGCACTGCACTGCCATGACCTTATACGTAATGGCTAC	1715
QY	547	HisArgGlnTyrlleSerProSerLeuValHisGlyGlnArgHisPro	561
DB	1716	CCAGTTTCACTACAGGCTCTCGTATATCCACTAATGAGAGACAGCA	1760

RESULT 6  
US-10-425-114-32180  
; Sequence 32180, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 32180  
; LENGTH: 2237  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73274D05\_FLI  
US-10-425-114-32180

Alignment Scores:  
Pred. No.: 3,69e-25 Length: 2237  
Score: 325.00 Matches: 151  
Percent Similarity: 41.80% Conservative: 109  
Best Local Similarity: 24.28% Mismatches: 251  
Query Match: 10.47% Indels: 111  
DB: 12 Gaps: 25

US-09-890-475-1 (1-609) x US-10-425-114-32180 (1-2237)

Qy 19 LeuLeuGlnArgHisGlnSer---GluGlnArgArgGluLeuProLysIleValGlu 37  
Db 353 CTGGTAGAAGGAGGAGGCTAGAGAAAGAAAGCTGAAGCTTGAGTTGATTGCA 412  
Qy 38 ThrGluSerThrSerMetAspIleThrIleGlnSerLysGlnProGlnPheLeuLys 57  
Db 413 GAGAAAGAGGCTAAAGTT-----TCTACAAAGAGCAGTCCTCAGTGAAT 457  
Qy 58 SerIleAspGluLeu-----AlaAlaPheSerValAlaValGluThrPheLysArgGln 75  
Db 458 CAGCTTCAGGAGCTAAGGATGCTGCTCTCTCTAGCAGAGGTTGCACAAAATAT 517  
Qy 76 PheAspAspLeuGlnLysHisIleGlu-----SerIleGluAsnAlaIleAspSer 92  
Db 518 AAGGTGGAGCTTGCTGAGATACCTTGAAGCAAGTGAAGCAAGCAAGCAAGCAAGTACC 577  
Qy 93 LysLeuGluSerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMet 112  
Db 578 TCATCATGACACAAATGATCTCGTGTTCAGAGGAGAT----- 619  
Qy 113 LeuSerProProArgAsnValSerValGluThrThrValThrValSerGlnProSer 132  
Db 620 -----ACCCCTGCTAGTGGTGGGTGGGAGGCA 646  
Qy 133 GlnGluIleValPro---GluThrSerAsnLysProGluGlyArgMetCysGluLeu 151  
Db 647 TCTGAAGCTTTACCTGTTGAGATACGCGGCCCTGATTGAAGCAACTTTGGAACAG 706  
Qy 152 MetCysSerLysGlyLeuArgLysThrIleThrAlaAsnIleSerAspGlnAlaLysLeu 171  
Db 707 ATGACACCAAGGCTTCTCAAAATTTCTTCAGAAAACCTGACAAAACCTGCTAGCTTT 766  
Qy 172 MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp 191  
Db 767 CGTGATGAACCTTCTGTTGCAATAAATGTCACACTGATCTGCACGCTTCGACTTAAT 826  
Qy 192 CysIleGlyLysPheTyrLeu-----GlnGlyArgArgAlaPheThrLysGluSerPro 209

Db 827 TCCCTGGAGGGTTTCTTCCACAGACCAAACTAATTCACCTGGGAGTAACATAATCC 886  
Qy 210 MetSerSerAlaArgGlnValSerLeuLeuIleLeuGluSerPhe-----LeuLeu 226  
Db 887 CTTGAGGTCCAGCGCAAGAGCTGCATTCTTTAATGGAAGCTATAGCACCTGCACTAGG 946  
Qy 227 MetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGlu 246  
Db 947 ATGAAGGAGGCTGGT---GGCGACAGCCCTTGGAGCTCTGAATTAAGCAAGCAAGCCAA 1003  
Qy 247 ThrAlaAlaValAlaValArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaGlu 266  
Db 1004 GAAATTGGCAGAGTGAAGAGTAAGTTA---GCTGAGATTGACCTGGATGCTTCTAAT 1060  
Qy 267 -----LysMetAspAlaArgGlyLeuLeuValAlaCysPheGlyValProSer 284  
Db 1061 GGCTATTCAATTGGAGGCACAGGCTTCTCGCAGCTTCTTACAACTTTTAATGTTGATT 1120  
Qy 285 AsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnGluIleAla 304  
Db 1121 GTGCTTGACGAAGATGAATATGCAAGATCGTAGTTGCTGTCTCTCGTCGCAAGCAGACT 1180  
Qy 305 GlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSer 324  
Db 1181 GCTGTATCCTGTCGCTCTCTTGGTCTTAATGAGAAAGTACCAGGTATCATTTGAGGAGTT 1240  
Qy 325 IleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp 344  
Db 1241 GTTAAGCGGCACAGCAAAATTTGATGCACTTCAATTCACAAAGCTTTTGGGCTTTCCGAG 1300  
Qy 345 LysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGlu 364  
Db 1301 ACATTTCCCTCCCTGACCTCTGCTGAAGACATATGTTGAGGAACAAAGATACAAATTGA 1360  
Qy 365 ArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGlnLeu 384  
Db 1361 AAT---AATGGCGATGCAACTGCGATCTCAATTAACGGATGACCCCAATCTAGGAGCTA 1417  
Qy 385 AlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLysGlu 404  
Db 1418 ATTGCTTTGAGGCTGTGATTAAGTGTATCGAAGAGCATTAAGCTT----- 1462  
Qy 405 LeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeu 424  
Db 1463 -----CAGAAGGAGTGTTCACCTGGACCTCTCCAGAGCGTGTCTTCTGAGCTG 1510  
Qy 425 AspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAla 444  
Db 1511 AAGCCCAAGGGTGAGAAAAGGCCATCAAGT-----GATGCT 1546  
Qy 445 LysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetProPro 464  
Db 1547 GGGCGTACTTAT-----GCAAGAAGCTTCAGGCCCTCGGCATTTCTATTCTTAGG 1597  
Qy 465 ValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArgAsp 484  
Db 1598 AGACCTCGCGGTCTCTGTTGGTTTCAGCTGCTCGCAGCGCTCCATTTCCG----- 1645  
Qy 485 AspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPhe 504  
Db 1646 -----GGCTATACCTGGCAGCGTCTCTCTGGG 1672  
Qy 505 ProHisArgSerArg-----ArgSerProGluTyrMetValPro 517  
Db 1673 CCGATGCCCTCCCGTGGTCTGCTCCAGTGCCTCCCGTGGCCCT---CTTCTGCCCT 1729  
Qy 518 LeuPro-----HisGlyGlyLeuGlyArgSerValTyr-----AlaTyr 530  
Db 1730 CTTCCTGATAGATATGGAGCTGCTGACCGGTACCATACACACCAGCCACCAACATAC 1789  
Qy 531 GluHisLeuAlaProAsnSerTyrSer-----Pro-GlyHisGlyHisArg 545

Db 1730 GACTCTGGTGCATTTTCCTCTTACAGTGAACCTTTTCAGTGGCCCAAAACCACTTCCAGTAC 1849  
Qy 545 GLeuHisArgGlnTyrSer-----ProSerLeuValHisG1 557  
Db 1850 ACTCAGAGGTAGTGGCAGCATCATCAATTCAGCCCTTAAAGTCGCTTATGGTGGC 1909  
Qy 557 yGlnArgHis-----ProLeuGlnTyrSerProProleHisGlyGlnGlnG1 573  
Db 1910 CCAGGAGCACCAGTGCAGAGCACCTTATCCAGGTTATG---CCAGCGCATCTGGCCCGCT 1966  
Qy 573 nLeuProTyrGlyGlnArgVal-----TyrArgHisSerProSerGluGluAr 590  
Db 1967 GCTTCCAGCAGTATGCAATATATATGATGGATCCGATATATCCCTCCCCAGCACCATAG 2026  
Qy 590 gTyr 591  
Db 2027 ATTC 2030  
RESULT 7  
US-10-425-114-5736  
; Sequence 5736, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 5736  
; LENGTH: 2300  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; FEATURE:  
US-10-425-114-5736  
Alignment Scores:  
Pred. No.: 3,858-25 Length: 2300  
Score: 325.00 Matches: 151  
Percent Similarity: 41.80% Conservative: 109  
Best Local Similarity: 24.28% Mismatches: 251  
Query Match: 10.47% Indels: 111  
Gaps: 25  
US-09-890-475-1 (1-609) x US-10-425-114-5736 (1-2300)  
Qy 19 LeuLeuGlnArgHisGlnSer---GluGlnArgArgArgGluLeuProLyLeValGlu 37  
Db 409 CTGTAGAGAGAGAGAGCATAGAGAAAGAGCTGAAGCTTGTAGTGTATTGCA 468  
Qy 38 ThrGluSerThrSerMetAspLeuThrGlyGlnSerLysGlnProGlnPheLeuLys 57  
Db 469 GAGAAAGAGGCTTAAAGTT-----TCTACAAAGAGCATGCTCCTCACTGAAT 513  
Qy 58 SerIleAspGluLeu-----AlaAlaPheSerValAlaValGluThrPheLysArgGln 75  
Db 514 GAGCTTCAGGAGCTAAGGATGCTGCTCTCTTCTAGCAGAGGTTTCGACAAAATAT 573  
Qy 76 PheAspLeuGlnLysHisIleGlu-----SerIleGluAsnAlaIleAspSer 92  
Db 574 AAGGTGGAGCTTGTCTGAGATACCTTGAAGCAAGTGAAGCAAGCAACAAAAGTAAATACC 633  
Qy 93 LysLeuGluSerAspGlyValValLeuAlaAlaArgAsnAsnPheHisGlnProMet 112  
Db 634 TCAATCAATGACAAATGATCTCTGCTTCAGAGGAGAAAT----- 675

Qy 113 LeuSerProProArgAsnAsnValSerValGluThrThrValThrValSerGlnProSer 132  
Db 676 -----ACCCCTGCTAGTGGGTGGGTGAGCA 702  
Qy 133 GlnGluIleValPro---GluThrSerAsnLysProGluGlyGlyArgMetCysGluLeu 151  
Db 703 TCTGAAGCTTTACTGTTGAGCTAAGCGGCCCTGTATTGAAGCAACTTTGTGAACAG 762  
Qy 152 MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeu 171  
Db 763 ATGCACACCAAGGCTTCTGAAATTTCTTTTTCAGAAAACTGTAGAAAACTTCTAGCTTT 822  
Qy 172 MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp 191  
Db 823 CGTGATGAACCTTTCTGTTGCACTAAATATGTCACACTGATCTCCGCGCTTCGTACTTAAT 882  
Qy 192 CysIleGlyLysPheTyrLeu-----GlnGlyArgArgAlaPheThrLysGluSerPro 209  
Db 883 TCCTGGAGGCTTTCTTCCACAGCACCAAACTAATTCACCTGGGAGTAAACATATGCC 942  
Qy 210 MetSerSerAlaArgGlnValSerLeuLeuLeuGluSerPhe-----LeuLeu 226  
Db 943 CTGAGGTCACGCGCAAGAGCTGCATCTTTTAAATGAAGCTATAGCACCTGCACCTAGG 1002  
Qy 227 MetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGlu 246  
Db 1003 ATGAAGGAGCTGGT---GGCGACGACCTTGGAGCTCTGAAATTAAGGAGCAAGCAAG 1059  
Qy 247 ThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaGlu 266  
Db 1060 GAAATTCGCGAAGAGTGAAGTAGTTA---GCTGAGATTGACCTGGATGCTTCTTAAT 1116  
Qy 267 -----LysMetAspAlaArgGlyLeuLeuLeuAlaCysPheGlyValProSer 284  
Db 1117 GGTATTCTATGAGGCACAGGCTTCTCTGACGCTCTTACAACITTTAATGTGATTCG 1176  
Qy 285 AsnPheArgSerThrAspLeuLeuAspLeuLeuLeuLeuValAlaCysPheGlyValProSer 304  
Db 1177 GTGCTTGGACAGATGAACTATGCAAGTCTGTGCTGTCTCTGCTGCGACGAGCT 1236  
Qy 305 GlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSer 324  
Db 1237 GCTGATCTCTGCTCTCTTGGCTTATGAGAAGTACCAGGTATCATTCAGGAGTTG 1296  
Qy 325 IleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPheGlyMetGluAsp 344  
Db 1297 GTTAAGCGGCACAGGCAAAATGATGAGTTCATTTTCATCAAGCTTTTGGGCTTCGAG 1356  
Qy 345 LysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGlu 364  
Db 1357 ACATTCCTCCCTGCACCTCTGCTGAAGCATATGTTGAGAACAAAGATACATTTGAA 1416  
Qy 365 ArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrLysGlnLeu 384  
Db 1417 AAT---AATGGCGATGCAACTGCGATCTCATTAACGATGACCCCAAACTCTAGGAGCTA 1473  
Qy 385 AlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAlaLysGlu 404  
Db 1474 ATTGCTTTGAGGCTGCTGATTAAGTGTATCGAAGCATTAAGCTT----- 1518  
Qy 405 LeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLysAspThrLeuGlnLeu 424  
Db 1519 -----CAGAAGGAGTGTTCACCTGGACCTCTCCAGAAGGCTGTCTTCTGAGCTG 1566  
Qy 425 AspLysGluMetGluLysAlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAla 444  
Db 1567 AAGCCCAAGGGTGAAGAAAGGCCATCAAGT-----GATGCT 1602  
Qy 445 LysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMetProPro 464  
Db 1603 GGGCGTACTTAT-----GCAAAGAGCTTCGAGGCGCTGCGCATTTCTTCTTAGG 1653  
Qy 465 ValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGlnArgAsp 484

```
Db 1654 AGACCTGCCGGTCTGTTGGTTCAGCTCGCAGGCTCCATTCCG-----1701
Qy 485 AspAspGlnAspGluIleSerAlaLeuValSerTyrLeuGlyProSerThrSerPhe 504
Db 1702 -----GGCTATACCTGGCAGCGTCTCCTGCG 1728
Qy 505 ProHisArgSerArg-----ArgSerProGluTyrMetValPro 517
Db 1729 CCGATGCCCTCCCGTGGTCTGCTCCAGTGCCTCCCTGCGCCCT---CTTCCTGCCCT 1785
Qy 518 LeuPro-----HisGlyGlyLeuGlyArgSerValTyr-----AlaTyr 530
Db 1786 CTTCCTGTAGTATAGGAGCTGCTACCGGTACCATTTACACACACCCAGCCATCCAGTAC 1845
Qy 531 GluHisLeuAlaProAsnSerTyrSer-----ProGlyHisGlyHisAr 545
Db 1846 GACTCTGGTGCATTTTCGTTTACAGTGAACCTTTTCAGTGCCGCCAAAACCATTCAGTAC 1905
Qy 545 GluHisArgGlnTyrSer-----ProSerLeuValHisG1 557
Db 1906 ACTCAGGGTCACTGGCAGCATCATCAATTCAAGCCCTATTAAGTCGTTTGGTGGC 1965
Qy 557 yGlnArgHis-----ProLeuGlnTyrSerProProIleHisGlyGlnGlnG1 573
Db 1966 CCAGGAGCACCACTCAGAACCACTATCCAGGTTATG---CCAGCGCATCTGGCGCGGCT 2022
Qy 573 nLeuProTyrGlyIleGlnArgVal-----TyrArgHisSerProSerGluGluAr 590
Db 2023 GCTTCCAGCAGTATGCAAAATTATATGGATCCGCTATATCGCTTCCCCCAACCATAG 2082
Qy 590 gTyr 591
Db 2083 ATTC 2086
```

## RESULT 8

```
US-10-425-114-32984
; Sequence 32984, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32984
; LENGTH: 1780
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17048G12_FLI
US-10-425-114-32984
```

```
Alignment Scores:
Pred No.: 1,32e-23 Length: 1780
Score: 309.50 Matches: 138
Percent Similarity: 41.27% Conservative: 89
Best local Similarity: 25.09% Mismatches: 223
Query Match: 9.97% Indels: 100
DB: 12 Gaps: 21
```

US-09-890-475-1 (1-609) x US-10-425-114-32984 (1-1780)

```
Qy 85 SerIleGluAsnAlaIleAspSerLysLeuGluSerAsnGlyValValLeuAlaAlaArg 104
Db 5 AGCAAGACAAAGGTAAGTACCTCAATCAATGACAAACATGCAATCTCGTGTTCAGAG 64
```

```
Qy 105 AsnAsnAsnPheHisGlnProMetLeuSerProProArgAsnAsnValSerValGluThr 124
Db 65 GAGAAAT-----ACC 73
Qy 125 ThrValThrValSerGlnProSerGlnGluIleValPro---GluThrSerAsnLysPro 143
Db 74 CTGCTAGTGGTTCGGCGGAGCATCTGAAGCTTACCTTGTGAGACTAAGCGCGCCCT 133
Qy 144 GluGlyGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAla 163
Db 134 GTATTGAAGCAACTATGTGAACAGATGACACCAAGGCTCTCTGAAATTTCTTTTCAGAA 193
Qy 164 AsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLysLeuAlaLys 183
Db 194 AACTGTAGAAAACCTTCTAGCTTTCGTGATGAACATTTCTGTGCACTAAATGTGCAACT 253
Qy 184 GluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeu-----GlnGlyArg 201
Db 254 GATCCTGCACGCTTCGTGCTTAATTCCTGGAGGGTTCTTCCACACAGACCAACTAAT 313
Qy 202 ArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeu 221
Db 314 TCACCTGGAGTAAACATAATGCCCTTGAGGTCACGCGCAAGAGCTGCATTTCTTTAATG 373
Qy 222 GluSerPhe-----LeuLeuMetProAspArgGlyLysGlyLysValLysIleGlu 238
Db 374 GAAGCTATAGCACTGCACCTAGGATGAAGAGGCTGGT---GGCGACGACCTTGGAGC 430
Qy 239 SerTyrIleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThr 258
Db 431 TCTGAAATTAAGGAGCAAGCAAGAAATTTGCCGAAGAGCTGGAAGTAGTAAAGTTA---GCT 487
Qy 259 GluGlyGlyLeuAlaAlaGlu-----LysMetAspAlaArgGlyLeuLeuLeuLeu 276
Db 488 GAGATTGACCTGATCTTCTTAATGGCTATTCATGGAGGACAGAGCTTCTTCGACGCTT 547
Qy 277 ValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuLeuArg 296
Db 548 CTTACAACTTTTAATGTGATTCGGTCTTGACGAAGATGAACATATGCAAGATCGTAGTT 607
Qy 297 MetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMet 316
Db 608 GCTGTCTCTCGTCAAGCAGAGCTGCTGATCTGCTCTCTCTTGGTCTTAATGAGAAA 667
Qy 317 ValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMet 336
Db 668 GTACCAAGTATCATTCAGGAGTGGTTAAGCGGCACAGCAAAATTTGATGAGTTCATTTC 727
Qy 337 ValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeu 356
Db 728 ATACAAAGCTTTTGGGCTTTTCAGAGACATTCCTCCCTGCACTCTGCTGAGACATATGTT 787
Qy 357 LysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPhe 376
Db 788 GAGGAACAAAGGATACAAATTGAAAT---AATGGCGATGCACTCGCATCTCAATTAACG 844
Qy 377 LysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThr 396
Db 845 GATGACCCCAAAATCTAGGAGCTAATTTGCTTTGAGGGCTGTGATTAAGTGTATCGAAGAG 904
Qy 397 HisLysLeuAspProAlaLysGluLeuProGlyTyrGlnIleLysGluGlnIleValSer 416
Db 905 CATAAGCTT-----CAGAAGAGTGTTCATCTTGGACCT 937
Qy 417 LeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSer 436
Db 938 CTCACAGAGCGTGTCTTCGTGAGCTGAAGCCCAAGGGTGAAGAAAGGTCATCAAGT----- 991
Qy 437 LeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArg 456
Db 992 -----GATGCTGGCGCTACTTAT-----GCAAGAAAGCGCTCGA 1024
```



QY 457 LeuSerProMetGluMetProProValThrSerSerSerTySerProIleTyArgAsp 476  
Db 1025 GSCCTGCAATTCATTTCTAGGAGAGCTGCGGTTCTGTGTTCAGCTGCTCGCAGG 1084  
QY 477 ArgSerPheProSerGlnArgAspAspGlnaspGluIleSerAlaLeuValSerSer 496  
Db 1085 COTCATTTCCG-----GGC 1099  
QY 497 TyrLeuGlyProSerThrSerPheProHisArgSerArg----- 509  
Db 1100 TATACCTGGCAGCGTCTCTCGCGATGCGCTCCGCTGCTGCTCGCTCCCTCC 1159  
QY 510 ArgSerProGluTyMetValProLeuPro-----HisGlyGlyLeuGlyArgSerVal 527  
Db 1160 COTGCCCT---CTTCTGCGCCCTCTCTCTGATAGATATGAGCTGCTGACCGGTACCAT 1216  
QY 528 Tyr-----AlaTyrlGluHisLeuAlaProAsnSerTyrSer----- 539  
Db 1217 TACACACACAGCCCGCAGCATACGACTCTGGTGCATTTTCTGTCTACAGTGAACCTTC 1276  
QY 540 -----Pro-GlyHisGlyHisArgLeuHisArgGlnTyrSer----- 551  
Db 1277 AGCGCCCAAAACCATTCAGTACACTCCAGGTCAGTGGCAGCATCATCAATCAAGC 1336  
QY 552 -----ProSerLeuValHisGlyGlnArgHis-----ProLeuGlnTyrSe 565  
Db 1337 CCCTATAAGTCGTTATGTTGGCCAGGAGCAGCAGTCAAGACCATTCAGTTAT 1396  
QY 565 rProProlleHisGlyGlnGlnLeuProTyrlGlyIleGlnArgVal-----Ty 582  
Db 1397 G---CCGCGCATCTGCGCGCGCTCTCCAGCAGCTATGCAAAATTATGGATCCGTA 1453  
QY 582 rArgHisSerProSerGluGluArgTyr 591  
Db 1454 TATCGCCCTCCCAACCATAGATTTC 1481

## RESULT 9

US-10-424-599-49586  
; Sequence 49586, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5323)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 49586  
; LENGTH: 2199  
; TYPE: DNA  
; ORGANISM: Glycine max  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_15784C.1  
US-10-424-599-49586

## Alignment Scores:

Prod. No.: 4,398-22 Length: 2199  
Score: 297.00 Matches: 134  
Percent Similarity: 38.65% Conservative: 89  
Best Local Similarity: 23.22% Mismatches: 256  
Query Match: 9.57% Indels: 98  
DB: 12 Gaps: 16

US-09-890-475-1 (1-609) x US-10-424-599-49586 (1-2199)

QY 46 ThrIleGlnSerLysGlnProGlnPheLeuLysSerIleAspGluLeuAlaPhe 65  
Db 81 ACGATGGGGTCCATCCCGATCCAGCGAGTTGACGATTGACTCAG----- 128

QY 66 SerValAlaValGluThrPheLysArgGln----- 75  
Db 129 ---CCGAGCTTCGAGAGTTCCAGGCCAAACTTCCCTCATGACGAGTGCACCTCCTC 185  
QY 76 PheAspAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGlu 95  
Db 186 TGAAGAGACTTCGACCACTTCTCTCCCTCCGAGCAGACCTCAACCAAAATCCGAA 245  
QY 96 SerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerPro 115  
Db 246 GCGCTCAAAACGCAAGATTGCGACCTCGACAACTACTTCCGACTCCTCGCTCCTC 305  
QY 116 ProArgAsnAsnValSerValGluThrThrValThrVal----- 128  
Db 306 GATCGCCGGAACCTCTCGACGCCAGCTCCAGATCGCCCTCGGACGCTCGACAG 365  
QY 129 -----SerGln 130  
Db 366 CGCGCAGCGCTGCTCTCCGCTCTCCACGCGCGGACGACATCATCAACAGCTCC 425  
QY 131 ProSerGlnGluIleValProGluThrSerAsnLysProGluGlyGlyArgMetCysGlu 150  
Db 426 CCGACGCGTGGAGTGCAGACACCCAGCCCTCATCTCAAACTCAAGTCTTCTGCTC 485  
QY 151 LeuMetCysSerLysGlyLeuArgLysTyrlleTyAlaAsnIleSerAspGlnAlaLys 170  
Db 486 CGCATGGACGCATTGGGTTCTTCGCTTCGTCAGCGCCCAAGAGAGCTCGACGC 545  
QY 171 LeuMetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeu 190  
Db 546 CTACGCGCGAGATGCTGTGCTCTGCGGAGTGGTTGATCCGCGCAAAATTCGTGCTG 605  
QY 191 AspCysIleGlyPheTyrlleGlnGlyArgArgAlaPheThrLysGluSerProMet 210  
Db 606 GAGCGATCTCGAGGTGTTTCGCTGACAGAGAGGGGCAAGCGCCGCGACGACTTG 665  
QY 211 SerSerAlaArgGlnValSerLeuLeuLeuGluSerPheLeu-----LeuMet 227  
Db 666 GCGTGGGCC-----TGCGTCTTGTCTCGGAGTGGTCTGATTCCGCTCGTGGAC 716  
QY 228 ProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThr 247  
Db 717 CCGTCATCGGAAATTCGAGGCTGTGGTGATCTCTATTGTGAAGGAGCAGCCAGGAG 776  
QY 248 AlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaGluLys 267  
Db 777 ATCGCAGACTTGGAGACACCGCTCGAGGAGCGGCTGGGTGGAACCTGAAGACC 836  
QY 268 MetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArg 287  
Db 837 CCGTGCCTCCACACTTCTTCGACGACGCTGTTACCTTCGGGATTGTCAAGAACGAGAC 896  
QY 288 SerThrAspLeuLeuAspLeuLeuArgMetSerGlySerAsnGluIleAlaGlyAlaLeu 307  
Db 897 TCGATTGTACCGGAGCTTGTATTGCTTCGCTTGGAGGAAACAGATGCCGAGCTC 956  
QY 308 LysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSerIleLysArg 327  
Db 957 GCGCTTTCGCTGCTCGCTCAGCAATGCTGATATGATTGAAGAGTTGATCAGCAA 1016  
QY 328 GlyMetHisIleGluAlaLeuGluMetValTyrlleThrPheGlyMetGluAspLysPheSer 347  
Db 1017 GGGCAGCAGCTTGATCGGTTTCACTTACATGAGTGGGTCTGTGTGAAGAGTTCCCT 1076  
QY 348 AlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLys 367  
Db 1077 CCGTTCCCTTGTGAAGTCTTTTCTCAAG-----GATGCTAAG 1115  
QY 368 ArgLysAlaGlnSerProLeu-----AlaPheLysGluAla 379  
Db 1116 AAAGTTGCGGCTTCTATTGGAAGATCTAAATGACGCGGAGCTCGCTACCTAGCT 1175  
QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399

```
Db 1176 GCAAGGAAGAGGAGCTGCACTCAGGCTGTGATTAAATGCAATGGAATCAAACTT 1235
Qy 400 AspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGluLys 419
Db 1236 GAG-----GATGAGTTCACACAGAGAGATCTGAGAGGAGCTTGACCACTGGAGAAG 1289
Qy 420 ---AspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSerLeuMet 438
Db 1290 GTGAAGACGGTGAAGACCGAAGAG-----1316
Qy 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnIleLysArgProArgLeuSer 458
Db 1317 -----AAACCGGTGCTGCTCTCCATTAAGAGAACTCGAGCAGC 1358
Qy 459 ProMetGlu-----MetProValThrSerSerSerTyrSerProIleTyrArg 475
Db 1359 AACAGCAATGAGGTGCAATGCCACGACCAAGCTGGGGGTTTGACTAATGCATATGTA 1418
Qy 476 AspArgSerPheProSerGln-----ArgAspAspAspGlnAspGluIleSer 491
Db 1419 TCA---TCCTTCCCTGCTGCTCTACATTTGTCAGGTCCTCATCAGGGCAATACCCA 1475
Qy 492 AlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSer 511
Db 1476 GCTGCTCTCCACCTTACCTTCCACCCACCATGATGTCAGCAGCAAGTCCCCGACA 1535
Qy 512 ProGluTyrMetVal-----ProLeuProHisGlyLeuGlyArgSerValTyrAla 529
Db 1536 AATCCTTATGCTGCTTATTCACCTGAGCCGACCCGCTATGTCAGGGTCTTACCCGGCA 1595
Qy 530 TyrGluHisLeuAlaProAsnSerTyrSerPro-----GlyHisGlyHisArgLeu 546
Db 1596 -----GTCCTCATGAATCTCTCTCTGATGATGGTGGCTATGGAATGTTTG 1643
Qy 547 HisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGln 563
Db 1644 GCTCCCACTTATCAGCAGGCTTACTACCGATAGAATGACAACCCCTTGAAG 1694

RESULT 10
US-10-424-599-95969
; Sequence 95969, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95969
; LENGTH: 1812
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57671C.1
US-10-424-599-95969

Alignment Scores:
Pred. No.: 6,95e-22 Length: 1812
Score: 294.00 Matches: 133
Percent Similarity: 38.69% Conservative: 86
Best Local Similarity: 23.50% Mismatch: 245
Query Match: 9.47% Indels: 102
DB: 12 Gaps: 20

US-09-890-475-1 (1-609) x US-10-424-599-95969 (1-1812)
Qy 62 LeuAlaAlaPheSerValAlaValGluThrPheLysArgGlnPheAspLeuGlnLys 81
```

```
Db 28 CTGAACACTCTTTCAGCGCAAGAAAGAGATTCTCAAGAAAGCTACTACGATCTCCAAAGT 87
Qy 82 HisIleGluSerIleGluAsnAlaIleAsp-----SerLysLeuGluSerAsnGly 98
Db 88 CACTCTCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 147
Qy 99 ValValLeu---AlaAlaArgAsnAsnAsnPheHis-----109
Db 148 ACCTCCCTCCACACCTCTCTCTCACACCGCTTCCACCTCTCTCTCTCTCTCTCTCTCTCTCT 207
Qy 110 -----GlnProMetLeuSerPro-----ProArgAsn 118
Db 208 CAACATATACCCACCATCTCTCTCCCAATACCTCAGCTTTCACCTCCACCTCCACCCAGAC 267
Qy 119 AsnValSerValGluThrValThrValSerGlnProSerGlnGluIleValProGlu 138
Db 268 CCATCATCGCAAAACGCGACCGCATTCGCCAAAATCCCTCGAACAATCTCTC-----321
Qy 139 ThrSerAsnLysProGluGlyArgMetCysGluLeuMetCysSerLysGlyLeuArg 158
Db 322 -----ACGCTCTGCAACAACATGCGAAGAAAGGGTTGAGG 357
Qy 159 LysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAla 178
Db 358 GACTAGCTGGGGACCATTTAAAGACAAGGCTGCAATCGAGGATACGCTTCGGAGTGCA 417
Qy 179 LeuLysLeuAlaLysGlu---ProAlaLysPheValLeuAspCysIleGlySerPheTyr 197
Db 418 CTCAGAGAGTGGCTCTGATGCTGCCGATCGATCTTCTCGATTCGTCGCGCGT-----471
Qy 198 LeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSer 217
Db 472 GTTGTGGGGCGGAATGTGGTGAAGATGACACGGAAATTCGGCTCGAGAAAGAGACTTGC 531
Qy 218 LeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysLe 237
Db 532 -----AGTTTTTGTTCAGCAGATTGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576
Qy 238 GluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMet 257
Db 577 AGTTTAAAGGAGAAATTAAAGCCCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 636
Qy 258 ThrGluGlyGlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuVal 277
Db 637 AGCGATGGT-----TGTGTGATGGTGTGTGGGGCCATGCGCTTCTTGCATTGCTG 687
Qy 278 AlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMet 297
Db 688 GCGGCTTATGGCTCTCTCTGAAATTGACCGTGCATGAGATTCTCACCCTCTCTCTTTATA 747
Qy 298 SerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetVal 317
Db 748 GCTGCTTCCAATGATGAGCTTCTGAGCTTACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Qy 318 SerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal 337
Db 808 CCAGGCTCTGTGCAAAAATTTATTCAGAGGAGCAAAACATATTCTGCGCTGTCAAGTTGT 867
Qy 338 TyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLys 357
Db 868 TTTGAGTTCAATCTGCTCATAGATTCCACGATTCACCATTTTGGAGGCTCATGTG---924
Qy 358 MetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLys 377
Db 925 -----AATGAGTCTCAGAAACTTGTAAAGACTTCTGAAGAGAGGAAAGTCACTTAGT 978
Qy 378 GluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHis 397
Db 979 GAGATCACGGCAGAGAAATCCACTCCACTGAAATCGGCGATTAAGGTTATTGAGGTGAT 1038
Qy 398 LysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeu 417
```

1039 AATCTTCAA-----TCTGAATATCCACCTGAAAGCCTTCAACAGCGTATAGACCAATTG 1092  
418 GluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaAArgSerLeuSerLeu 437  
1093 -----ATGAGCATATAGGCA---AATGTAATAATAT 1119  
438 MetGluGluAlaAlaLeuAlaLysArgMetTyrAenGlnGln-----IleLys 453  
1120 GCTGATCGGCTTTTCTGCAAGCCTCTCCACATCAGACGCAACAAGTGAATCAAG 1179  
454 ArgProArgLeuSerProMetGluMetProProValThrSerSerSerTyrSerProile 473  
1180 CTTCTCGAATGTC-----GAGCCAGTTGGTTCTGCTTCTGCTCTGAACAGT 1227  
474 TyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGluLysSerAlaLeu 493  
1228 GCCAGTGGTGGCCAGCTCAACCGTTCACTACCAACAACCTCATTTCCAGTCATCAGGTTG 1287  
494 ValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSerArgSerProGlu 513  
1288 TTGCTGGAACTATTAATCCGTACATGAACTTGCCA----- 1323  
514 TyrMetValProLeuProHisGlyGlyLeuGlyArg-----SerValTyrAlaTyrGlu 531  
1324 -----ACCATCCATATGCGCATGAAGGCTCAAAACCCATCCATCCCGCTTATACA 1374  
532 HisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSer 551  
1375 GCGCTTCAACCTGAGCGCTTATGCTCTGAT-----GGTGTCCCAATGGGT 1419  
552 ProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerProProileHisGlyGln 571  
1420 CTTAGTGCAACCGTGGTCAAGTGGTCTCTT-----CCATTTTCATCAGAA 1467  
572 GlnGlnLeuPro-----TyrGlyGlnGln 579  
1468 CCACTTATGCGCTCGGTATTATGATATGATGCTGCTGCTGTTATGGTGTATGCTGCA 1527  
580 ArgValTyrArgHisSer 585  
1528 CATTATTACCAACATCT 1545

## RESULT 11

US-10-424-599-49595  
; Sequence 49595, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 49595  
; LENGTH: 2393  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_15792C.1  
US-10-424-599-49595

## Alignment Scores:

Pred. No.:	1,78e-21	Length:	2393
Score:	292.00	Matches:	127
Percent Similarity:	41.12%	Conservative:	86
Best Local Similarity:	24.52%	Mismatches:	240
Query Match:	9.41%	Indels:	66
DB:	12	Gaps:	15

US-09-890-475-1 (1-609) x US-10-424-599-49595 (1-2393)  
QY 69 ValGluThrPheLysArgGlnPheAspLeuGlnLysHisIleGluSerIleGluAsn 88  
DB 285 ATTACACCTCGACAACTCCACTCGGACTCCCTCCGCTCCTCGACACCGCAACCC 344  
QY 89 AlaIleAspSerLysLeuGluSerAenGlyValValLeuAlaAlaArgAsnAsnPhe 108  
DB 345 TCTCTCGAGCCACTCTCCAG-----ATCGCCCTCGCAGCTCGACGACG 389  
QY 109 HisGln-ProMetLeuSerProProArgAsnAsnValSerValGluThrThrValThrVa 128  
DB 390 CGCGGACGCTGCTCTCTCCGCTCTCTCC-----ACGACGCTGACGAC 434  
QY 128 lSerGlnProSerGlnGluLeuValProGluThrSerAenLysProGluGlyGlyArgMe 148  
DB 435 ACCTCCCGGACGCGGAGGTTCGATGACACCGCGGCTCGCTCAAGCTCAAGTCC-TT 493  
QY 148 tCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGl 168  
DB 494 CTGCTCCGATGAGCGGCTTGGCTTCTCGCTTCGTCAGCGCCCAAGAGGAGCT 553  
QY 168 nAlaLysLeuMetGluGluLeuProSerAlaLeuLysLeuAlaLysGluProAlaLysPh 188  
DB 554 GAGCGCTTACGCGCGGAGATCGGCTGGCGGAGTGGCTCGATCCGCGGCAAAAT 613  
QY 188 eValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArgAlaPheThrLysGluSe 208  
DB 614 CGTGCCTGGAGGCGATCTCGAGGCTTCCGCTCGACAGAGAGGG-----GAGAA 664  
QY 208 rProMetSerSerAlaArgGlnValSerLeuLeuLeuLeuGluSerPheLeu----- 225  
DB 665 GCGCGCCACGACTTGGCTGGCGCTCGCTGCTGCTGGAGTCTTAATTCGCTGCT 724  
QY 226 -LeuMetProAspArgGlyLysGlyLysValLysLysLeuSerTrpIleLysAspGluAl 245  
DB 725 CGTCGACCCGCTCATCGAAATTCGAGGTGTGGTTACCTTACCTACCGTGAAGAGGATGC 784  
QY 245 aGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGlyGlyLeuAlaAla 265  
DB 785 CACGGAGATCGCGGAGACTTGGAGAGCAGCTGGAGGATCGTGGCGGTGGAACCT 844  
QY 265 aGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyValProSerAs 285  
DB 845 GAAGACACCTGACCTCCACACTTCTTGACGACGCTGTGTACCTTCGGGATGTCAAGAA 904  
QY 285 nPheArgSerThrAspLeuLeuAspLeuLeuArgMetSerGlySerAsnGluLeuAlaGl 305  
DB 905 CGATGACTCGATTTGATCGGAGCTTGTATTGCTTCGCTTGGAGGAAACAGATGCC 964  
QY 305 yAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValGluSerSerIl 325  
DB 965 GAAGCTCGCGCTTTCGCTTGGCTCGCTCAGCAATGCTGATATGATTGAAGAGTGTAT 1024  
QY 325 eLysArgGlyMetHisIleGluAlaLeuMetValTyrThrPheGlyMetGluAspLys 345  
DB 1025 CAGCAAAAGGCGACGCTTGGTTCCTTACCTTATGAGTGGGTCTGCTGGTGAANA 1084  
QY 345 sPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSerPheGluAr 365  
DB 1085 GTTCCCTCTGTTCCCTTGTGAGTCTTTTCTCAG-----GA 1123  
QY 365 gAlaLysArgLysAlaGlnSerProLeu-----AlaPheLys 377  
DB 1124 TGCTAAGAAAGTTGCGGCTTCTATTTTGAAGATCCTAATAATGACGCGGAGCTCGGTA 1183  
QY 377 sGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHi 397  
DB 1184 CTTAGCTGCAGGAAAGAGCAGCTTCGACTCAGGCTGTGATTAAATGATTGAAGAAATA 1243  
QY 397 sLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLe 417  
DB 1244 CAAACTTGAG-----GATGAGTTCCTCCAGAAATCTGAAGAGCGAGCTTGACCAATT 1297

```
Qy 417 uGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaAlaArgSerLeuSerLe 437
Db 1298 AGAAG-----GTGAAGATGGTGAAGACCGAGAAAGGAAACCAAGTGGCAGT 1345
Qy 437 uMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLe 457
Db 1346 TCCTGCCAAT-----AAGAGAACTAGAGC 1369
Qy 457 user-----PrometGluMetProProValThrSerSerSerTyrSerPr 472
Db 1370 AAGCAACGGCAATGGAGTCCA-----ATGCCACCGACCAAGCTGGGGTTGACTAA 1423
Qy 472 oLleTyrArgAspArgSerPheProSerGln-----ArgAspAspAspGlnAs 488
Db 1424 TGCATATGTATCA---TCTTTCCTGCTGCTCTTACATTTGTCCAGGTCCTCATCACACGG 1480
Qy 488 pGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThrSerPheProHisArgSe 508
Db 1481 GCATATCCAGCTCTCTTCCACCATACCTTCCCCACCCACACATGTACGGCAGCAGAAG 1540
Qy 508 rArgArgSerProGluTyrMetVal-----ProLeuProHisGlyGlyLeuGlyArgSe 526
Db 1541 TCCCCAGCAAAATCTTATGCTTATTCACCCGAGCGGCACCACTATTGCAGGGTC 1600
Qy 526 rValTyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLe 546
Db 1601 TTACCCGGCAGTCCCTGAACTATCTCATGCAATATGGC---GGCTATGGAATGTTTT 1657
Qy 546 uHisArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGln 563
Db 1658 GGCTCCCACTTATCAGCAGGCTTACTACCGATAGAAATGCAACCTTTGAAG 1709
```

## RESULT 12

```
US-10-425-114-34508
; Sequence 34508, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34508
; LENGTH: 1772
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO1727806_FLI
US-10-425-114-34508
```

```
Alignment Scores:
Pred No.: 9,32e-19 Length: 1772
Score: 265.50 Matches: 114
Percent Similarity: 39.74% Conservative: 68
Best Local Similarity: 24.89% Mismatches: 206
Query Match: 8.55% Indels: 71
DB: 12 Gaps: 12
```

US-09-890-475-1 (1-609) x US-10-425-114-34508 (1-1772)

```
Qy 160 TyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeu 179
Db 18 TTCGTGTGGCCCGCCGCAAGAGGAGCGCGTGTGGGGCCGAGATGCGCGCGGCGTC 77
Qy 180 LysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGln 199
```

```
Db 78 AAGCTCTCGCTGGATCCCGCCCAAGTTCGTATGGACGCGCTGGCGAGCGTCTTCCCGCTC 137
Qy 200 GlyArgAlaAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeu 219
Db 138 GACCGCCGCGAGGTCGGGAACCCCGCCACCTGGCTGGCC-----TGCGTGTCTC 188
Qy 220 IleLeuGluSerPheLeu-----LeuMetProAspArgGlyLysGlyLys 234
Db 189 ATCTCTGAGGCGCGCTGCTCCCGCTCCGACACGCGACATCGGCCCGCGCGC 248
Qy 235 ValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAlaTatPArgLys 254
Db 249 CCGCTCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 308
Qy 255 ArgLeuMetThrGluGlyGlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeu 274
Db 309 GCGCGCGAGAAAGAGGGGGGCTCGAGGGCGCGAAGCGCGCGCGCGCGCGCGCGCTTCTG 368
Qy 275 LeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeu 294
Db 369 CAGCTCGTGGCGCGCGCTTTCGCGCTCCCGAGAGG---GAGGACAGCGCGCTGTACCGCAGG 425
Qy 295 IleArgMetSerGlySer---AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeu 313
Db 426 ATTGCTGTCAGCTTCTCTGCGCGCGCGCGCGCGCGCGCTCGCGCTCACCTCGGACTC 485
Qy 314 ValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIleGluAla 333
Db 486 GAGGAACAAATGGCGGATATCATTTGAGGAATGTTGTAAGAGGACGACGCTTGATGCT 545
Qy 334 LeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThr 353
Db 546 GTAATTTGCTTATCAGCTGGGCTTCAGAGAGAGTTCCACCAAGTTCCTTTTGAAG 605
Qy 354 SerPheLeuLysMetSerLysGlySerPheGluArgAla-----LysArgLysAlaGln 371
Db 606 TCCTACCTGGAAGACTCTAAGAAGACATCAACCGCTGCTTCAGATAATTAAGCACTAGC 665
Qy 372 SerProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerValMet 391
Db 666 ACGGCGCAATCAGGAGCAATGTGAACAAGAAAGAACAGTCTGCACCTCGAGTGTGATA 725
Qy 392 GlnCysMetGluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLys 411
Db 726 AAGTGTGTGAGGATCGTAAACTAGAGCTGAGTTTCCACTGGAGGGT-----773
Qy 412 GluGlnIleValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGlyLys 431
Db 774 -----CTTCGGAAGCAACTTGAAGAACTAGAGAAAGCAAGCAAGCCGAGAG 818
Qy 432 AlaArgSerLeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGln 451
Db 819 AAGAAGGAGCATCA-----AGCGCTACAGTGGCGGCGAGCGAGCGGCCCAACA 866
Qy 452 IleLysArgProArgLeuSer-----Promet-----460
Db 867 ACCAAGCGCATCCGTGCGAGCAGCTGGAGGCGCAATGCTCTCGCCAGGCGAGTGTCTC 926
Qy 461 -----GluMetProValThrSerSerSerTyrSerProIle 473
Db 927 ACTAACAATGTTGGTGTCTTCTCTCCAGCTCCACCATGACCATGCTCTCCCTCC 986
Qy 474 TyrArgAspArgSerPheProSerGlnArgAspAspGlnAspGluIleSerAlaLeu 493
Db 987 CACGGGTATACGCCAACACCTCCCGCTCCGACCTCATACGCAATATCTCCCATCC 1046
Qy 494 ValSerSerTyrLeuGlyProSerThrSerPheProHisArgSerArgSerProGlu 513
Db 1047 CAGCGCTCATACGCCACAGCTCTCCCATCCCATGTCATCATACGACACCTCCCT---1103
Qy 514 TyrMetValProLeuProHisGlyLeuGlyArgSerValTyrAlaTyrGluHisLeu 533
```

Db	292	CGCGCGTGGATT	CGCGGGAGAAAGCTAGAGACTGGTGATCGGAGATACCGGAGAACTC	351
Qy	144	GlulGly	-----ArgMetCysGluIeuMetCysSerLys	155
Db	352	GACGACGGCAGCGGCTTTATCGCGCTGAAGTCGCTTTGTTGAAGATGGACGCAAGA	411	
Qy	156	GlyIeuArgLysTyrlTyAlaAsnIleSerAspGlnAlaLysIeuMetGluGluIle	175	
Db	412	GGATTTTGGGGTTTGATTTGCGAGGAAGAAGATTTGGAGATATCTCCGGTCACAGATT	471	
Qy	176	ProSerAlaIeuLysIeuAlaLysGluProAlaLysPheValIeuaspCysIleGlyLys	195	
Db	472	CCGTGTGCGGTTGGTAGATTGTGGATCCGCCGAAGTTAGTACTTGAAGCTGATCTCAG	531	

```

RESULT 13
US-09-938-842A-917
; Sequence 917, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME. AND METHODS OF USE

```

Qy	253	ArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaAlaGluLysMetAspAlaArgGly	272
	:	:	:
Db	709	AAGGCTAGCTTGAAGAGAGAGAGGGAATTGAGAATGTGAAAACACCTGATTTTCATACG	768
Qy	273	LeuLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeuLeu	292
	:	:	:
Db	769	TTTCTGCAACATCTGTGCATTTTGGGATTGTGAGCAG-----GATGATCTA	816
Qy	293	AspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGluPhe	312
	:	:	:
Db	817	GCTCTTTATAGG-----AAACTTGTGTGTTGGTTCAGCTTGGCGTTAAACAGATG	864
Qy	313	-----LeuValProMetValSerGlyIleValGluSerSer	324
	:	:	:

Alignment Scores:	
Pred. No.:	1,93e-15
Score:	234.50
Percent Similarity:	34.38%
Best Local Similarity:	1.934%
Query Match:	7.55%
DB:	9
	Gaps:
	Indels:
	Mismatches:
	Conservative:
	Matches:
	Length:

365	QY	-----ArgAlaLysArgIysAlaGlnSerProLeu	374
1045	DB	TTAATCACTGATGATTCACAAATTCGGCCGATCTCGGGATCTGTTTCTTCTGCACCTC	1104
375	QY	AlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMet	394
1105	DB	GCATTAC	CAATGATG
395	QY	GluThrHisLysLeuAspProAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIle	414
1120	DB	TGCTAGAGAAGACCAAAACCGAAGAGAAAACACGCC	1161
			-----GTG
415	QY	ValSerLeuGluLysAspThrLeuGlnLeuAspLysGluMetGluGluLysAlaArgSer	434
1162	DB	ATTCCTCCCAACACAGAGAAC	1182
435	QY	LeuSerLeuMetGluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnIleLysArg	454
1183	DB	-----CGACCAAGCTACACGGT	1200
455	QY	ProArgLeuSerProMetGluMetProProValThrSerSerSerTyrSerProIleTyr	474
1201	DB	-----CCAATGCCACCAAGCAGCGCGCTATCACA	1233





```

Qy 555 ValHisGly-----GlnArgHis-Pr 561
Db   ::: :|||
Qy 935 ATCAG-GGAGCCATATGTTTACCCAGCTGAGGAGGTAGCCACCGTCAACGTCGGCATTC 993
Db   ::: :|||
Qy 561 oLeuGlnTyrSerProProlleHisGlyGlnGlnGlnLeuProTyrGly 577
Db   ::: :|||
Qy 994 CATGGCCTACTCCACCCCGCCCATGAGCTACCCCTGCTGCCCCCTACGCG 1042

```

Search completed: February 28, 2004, 07:13:50  
Job time : 558 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 28, 2004, 00:18:24 ; Search time 3455 Seconds  
(without alignments)  
5263.696 Million cell updates/sec

Title: US-09-890-475-1  
Perfect score: 3104  
Sequence: 1 MSNPPTVAQPTTANPLL.....RYLGSLNQSRPSNSLDPK 609

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delopt 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/cgn2\_1/USPFO\_spool\_p2n.model -DEV=xlp  
-MCPDEL-frame+ p2n.model -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USPR=US09890475 @CGN 1 1 3549 @runat 24022004 135312 5304 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	742.5	23.9	704	28	BH491272	BH491272 BOHH16TR
C 2	696	22.4	706	28	BZ035667	BZ035667 Oe83610.
C 3	687.5	22.1	781	28	BH479860	BH479860 BOGRJ75TF
C 4	663	21.4	552	9	AV541057	AV541057 AV541057
C 5	612	19.7	750	28	BH470959	BH470959 BOHLJ19TR
C 6	580	18.7	931	14	CK252208	CK252208 EST738845
C 7	556.5	17.9	943	14	CK252381	CK252381 EST736018
C 8	550.5	17.7	962	14	CK286517	CK286517 EST749239
C 9	543	17.5	635	28	B77833	B77833 T29E23TR TA
C 10	535.5	17.3	905	14	CK243005	CK243005 EST726642
C 11	528	17.0	913	14	CK294055	CK294055 EST756769
C 12	525.5	16.9	945	14	CK288931	CK288931 EST751653
C 13	523.5	16.9	684	14	CD836567	CD836567 BN45.049C
C 14	519.5	16.7	883	14	CK249078	CK249078 EST732715
C 15	518.5	16.7	904	14	CK243006	CK243006 EST726643
C 16	502.5	16.2	945	14	CK246861	CK246861 EST730498
C 17	494	15.9	840	14	CK286207	CK286207 EST748929
C 18	491.5	15.8	971	9	AJ568038	AJ568038 AJ568038
C 19	487	15.7	520	28	BZ512352	BZ512352 BOMSL50TF
C 20	486.5	15.7	785	13	BQ115269	BQ115269 EST60845
C 21	479	15.4	827	14	CK259125	CK259125 EST742762
C 22	467.5	15.1	557	10	BF050953	BF050953 EST436111
C 23	461	14.9	771	14	CK294226	CK294226 EST756940
C 24	456.5	14.7	703	14	CK252362	CK252362 EST735999
C 25	440	14.2	771	12	BI176101	BI176101 EST517184
C 26	435.5	14.0	688	28	BH984132	BH984132 Oe42406
C 27	428	13.8	850	14	CK257335	CK257335 EST740972
C 28	423.5	13.6	671	12	BG890481	BG890481 EST516332
C 29	420.5	13.5	872	14	CK248078	CK248078 EST737175
C 30	392.5	12.6	896	14	CK258193	CK258193 EST741830
C 31	386.5	12.5	961	14	CK258053	CK258053 EST741690
C 32	385	12.4	685	10	AW738032	AW738032 EST739459
C 33	373	12.0	651	10	AW738612	AW738612 EST7340039
C 34	371	12.0	637	13	BQ118418	BQ118418 EST603994
C 35	353.5	11.4	540	10	BE471538	BE471538 EST416391
C 36	352.5	11.4	899	14	CK288916	CK288916 EST751638
C 37	352	11.3	500	13	BQ511650	BQ511650 EST619065
C 38	321.5	10.4	2069	11	AY109393	AY109393 Zea mays
C 39	301	9.7	877	14	CK263597	CK263597 EST709675
C 40	295	9.5	544	10	BF053940	BF053940 EST439170
C 41	291	9.4	695	14	CK273406	CK273406 EST719484
C 42	287	9.2	861	29	CG962836	CG962836 MBH9191TR
C 43	283.5	9.1	964	14	CK275053	CK275053 EST721131
C 44	279.5	9.0	430	12	BG097426	BG097426 EST462041
C 45	277.5	8.9	765	14	CB970934	CB970934 CAB10004

ALIGNMENTS

RESULT 1  
BH491272/c  
LOCUS BOHH16TR BOHB Brassica oleracea genomic clone BOHH16, genomic survey sequence.  
DEFINITION BH491272  
ACCESSION BH491272  
VERSION BH491272.1 GI:17699376  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 704)  
 AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.  
 TITLE Whole genome shotgun sequencing of Brassica oleracea  
 JOURNAL Unpublished (2001)  
 COMMENT Other GSSs: BOHBH16TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 1. 704  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TOL000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOHBH16"  
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.98e-58 Length: 704  
 Score: 742.50 Matches: 157  
 Percent Similarity: 74.80% Conservative: 27  
 Best Local Similarity: 63.82% Mismatches: 35  
 Query Match: 23.92% Indels: 28  
 Gaps: 7  
 US-09-890-475-1 (1-609) x BH491272 (1-704)  
 QY 372 SerProLeuAlaPhe-----LysGluAlaAlaThrLysGlnLeuAlaVal 386  
 Db 701 TCGTGTGGCTTTTATTTCTCAACAGAAAAGCTGCGAAGAAAGCAGTTCGATGCG 642  
 QY 387 LeuSerSerValMetGlnCysMetGlnThrHisLysLeuAspProAlaLysGluLeuPro 406  
 Db 641 TTATCATCAGTGAAGATTGGAGCTCACAGTTAGACCAGTGAAGAAAGTACCA 582  
 QY 407 GlyTrpGlnLeuLysGluGlnLeuValSerLeuGluLysAspThrLeuGlnLeuAspLys 426  
 Db 581 GGGTGGCAGATCCAGAGCAATGCGAAGCTTGAGAAAGAGATTGTTTCAGCTCGACAA 522  
 QY 427 GluMetGlu-----GluLysAlaArgSerLeuSerLeuMet 438  
 Db 521 CAGATGGGAAGACGAGATCCATCAGTCGAATGGAGGAGCGCATCCATCAGTCTAAGG 462  
 QY 439 GluGluAlaAlaLeuAlaLysArgMetTyrAsnGlnGlnLeuLysArgProArgLeuSer 458  
 Db 461 GAGGAAGCGGCAATTAGCGAGAGATTGTATTAACCAACAGATGAAAGCTCCAAAGTTGTCA 402  
 QY 459 ProMetGluMetProProValThrSerSerSerTyrSerProLysTyrArgAspArgSer 478  
 Db 401 GAAGGGAATGCCACCAACAGCTTCTTATCTCTTATCTCTATGATGATCCGCGACCAAGC 342  
 QY 479 PheProSerGlnArgAspAspGlnAspGluLysSerAlaLeuValSerSerTyrLeu 498  
 Db 341 TTCCCTAGTCACAGAGAGGAGATGCAGATGAATATCAGCTCTTGTGAGTAGTACCTC 282  
 QY 499 GlyProSerThrSerPheProHisArgSer-----ArgArgSerProGluTyrMetVal 516  
 Db 281 GGCCCATCAGCAGGTTTCTCATCGTCAAGTCTTCAGGAGATCCCTCGAATATATGTT 222  
 QY 517 ProLeuProHisGlyGluLeuGlyArgSerValTyrAlaTyrGluHisLeuAlaProAsn 536  
 Db 221 -----CACCTGGTGGGTAGAGAAAGTGTCTCTGCGTATGATCATCAGCTCCCAAT 168  
 QY 537 SerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSerProSerLeuValHis 556

Db 167 TCTTATTCGCG-----GTTTCAAGAGTACTCTCCA-----GTTTCAAC 129  
 QY 557 GlyGlnArgHisProLeuGlnTyrSerProLysHisGlyGlnGlnLeuProTyr 576  
 Db 128 GGACAGAGACTTCCCTCAAGAGTACTCTCTCCAGTTTCATGGCAACACCAATGCCATAT 69  
 QY 577 GlyLeuGlnArgValTyrArgHisSerProSerGluGlnArgTyrLeuGlyLeuSerAsn 596  
 Db 68 GGTCTA-----TACAGACATTCCCATCTGT-GAAGATACTTGGCTTTGTCCAAT 19  
 QY 597 GlnArgSerProArgSer 602  
 Db 18 CACAGACTCTCTCGTAAC 1  
 RESULT 2  
 BZ035667/c  
 LOCUS oed83e10.g1 B.oleracea002 Brassica oleracea genomic, genomic survey  
 DEFINITION sequence.  
 ACCESSION BZ035667  
 VERSION BZ035667.1 GI:23612705  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 REFERENCE 1 (bases 1 to 706)  
 AUTHORS Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,  
 Nash, W., Rabinowicz, P.D. and Wilson, R.K.  
 TITLE Whole genome shotgun reads from Brassica oleracea  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@watson.wustl.edu  
 Plate: oed83 row: e column: 10  
 Seq primer: -28RPrOT reverse  
 Class: shotgun  
 High quality sequence start: 17  
 High quality sequence stop: 551.  
 FEATURES  
 1. 706  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.oleracea002"  
 /note="Vector: pOTW13; Whole genome shotgun library from  
 flowering buds. DNA was purified from a crude nuclear  
 prep using Brassica oleracea TOL000DH3 buds provided by  
 Thomas Osborn at the University of Wisconsin. Genomic  
 DNA was provided by Pablo Rabinowicz (CSHL) and the  
 shotgun library prepared at Washington University Genome  
 Sequencing Center."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.18e-53 Length: 706  
 Score: 696.00 Matches: 146  
 Percent Similarity: 75.11% Conservative: 26  
 Best Local Similarity: 63.76% Mismatches: 35  
 Query Match: 22.42% Indels: 23  
 Gaps: 6  
 US-09-890-475-1 (1-609) x BZ035667 (1-706)  
 QY 377 LysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThr 396  
 Db 657 AAAACTGCCAAGCAAGCAAGCAGTGGTATCATCATCATGATGAAGTGT-TT-GAAGCT 599  
 QY 397 HisLysLeuAspProAlaLysGluLeuProGlyTrpGlnLeuLysGluGlnLeuValSer 416



DEFINITION AV541057 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
 AV541057 cDNA clone R2159d09F 3', mRNA sequence.  
 ACCESSION AV541057 GI:8702815  
 VERSION AV541057.1 GI:8702815  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (chale cress)  
 ORGANISM Arabidopsis thaliana  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 552)  
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
 of 12,028 non-redundant expressed sequence tags from normalized and  
 size-selected cDNA libraries  
 JOURNAL DNA Res. 7 (3), 175-180 (2000)  
 MEDLINE 20363093  
 PUBMED 10907847  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.  
 Location/Qualifiers  
 1..552  
 /organism="Arabidopsis thaliana"  
 /mol\_type="RNA"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="R2159d09F"  
 /tissue\_type="roots"  
 /clone\_lib="Arabidopsis thaliana roots Columbia"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:  
 XhoI"

## ORIGIN

Alignment Scores:  
 Pred. No.: 8,8e-51 Length: 552  
 Score: 663.00 Matches: 122  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 21.36% Indels: 0  
 DB: 9 Gaps: 0

US-09-890-475-1 (1-609) x AV541057 (1-552)  
 QY 488 AspGluIleSerAlaLeuValSerTyrLeuGlyProSerThrSerProHisArg 507  
 DB 551 GATGAATATAGCTCTTGAGTAGTACTCGCCCGCTCAACATCTTTCTCATCGC 492  
 QY 508 SerArgSerProGluTyrMetValProLeuProHisGlyGlyLeuGlyArgSerVal 527  
 DB 491 TCAGAAGATCCCGGAATATATGTTCCACTTCCACATGTTGGGTTAGGAAGAAGTGA 432  
 QY 528 TyrAlaTyrGluHisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHis 547  
 DB 431 TATGCATATGAACATCTGCGCCCAATATCATTTCTCCAGGTCAGGCATAGACTTCAT 372  
 QY 548 ArgGlnTyrSerProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerProPro 567  
 DB 371 CGACGACTCTCCGCTTTGGTTTCACGACAGACATCCACTACATGACTCTCTCTCCA 312  
 QY 568 IleHisGlyGlnGlnGlnLeuProTyrGlyIleGlnArgValTyrArgHisSerProSer 587  
 DB 311 ATTCATGGACACACACAGTTACCATGTATACAAAGGGTTACAGACATTCACCATCT 252  
 QY 588 GluGluArgTyrLeuGlyLeuSerAsnGlnArgSerProArgSerAsnSerSerLeuAsp 607  
 DB 251 GAAGAAAGATATTGGGTTTATCCAAATCAAGGTTCTCTCGCAGTAACTCATCTAGAC 192  
 QY 608 ProLys 609  
 DB 191 CCAAAA 186

RESULT 5  
 BH470959/c  
 LOCUS  
 DEFINITION BOHLJ19TR BOHL Brassica oleracea genomic clone BOHLJ19, genomic  
 survey sequence.  
 ACCESSION BH470959 GI:17679070  
 VERSION BH470959.1 GI:17679070  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 750)  
 Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.  
 Whole genome shotgun sequencing of Brassica oleracea  
 JOURNAL Unpublished (2001)  
 COMMENT Other GSSs: BOHLJ19TF  
 Contact: Chris Town

## TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..750

/organism="Brassica oleracea"

/mol\_type="genomic DNA"

/strain="TO1000DH3"

/db\_xref="taxon:3712"

/clone="BOHLJ19"

/clone\_lib="BOHLJ19"

/note="Vector: pHOS1; Site\_1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHOS1 using BstXI linkers"

## ORIGIN

Alignment Scores:

Pred. No.: 7,48e-46 Length: 750

Score: 612.00 Matches: 135

Percent Similarity: 70.22% Conservative: 23

Best Local Similarity: 60.00% Mismatches: 34

Query Match: 19.72% Indels: 35

DB: 28 Gaps: 2

US-09-890-475-1 (1-609) x BH470959 (1-750)

QY 313 LeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMethHisIleGlu 332

DB 666 ATTATCTATCGTCTAGGTATAGTTGATTTAAGTATCAAGCGTGAAGCATATTGAA 607

QY 333 AlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeu 352

DB 606 GCATTTGGATGATTTATACCTT-GGATAGAGATAGTTTTCGGCTTCTCGCTTCTA 548

QY 353 ThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGlnSer 372

DB 547 ACTTCATCTTAAAGTAGAGAGAGTCAATTTAGAGGGCAACACAGAAAGCTCAAGCA 488

QY 373 ProLeuAlaPheLys----- 377

DB 487 CCGATAGCATTTCT-ATGGCCCTTCTTAGCTTCTTAGCTCTCTCTTTTCTTCT 429

QY 378 ----- 382

DB 428 TTTCTGATCTTACCATTGTTGCCCTTTGTTCTCTTAACAGAAAGAGCCCAACCAAG 369

QY 383 GlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspProAla 402

DB 368 TTTTACGTCCTTGTATCATGATGAAGTGTTCGGAGGCTCACAACTTAGACCCAGAG 309

QY 403 LysGluLeuProGlyTyrGlnIleLysGluIleValSerLeuGluLysAspThrLeu 422  
 Db 308 AAGAAGTACAGGATGCGAGATCAAGACAAATGATTAAAGTTGGAGAAAGACATTCCT 249  
 QY 423 GlnLeuAspLysGluMetGluGluLysAlaArgSerLeuSerLeuMetGluGluAlaA 442  
 Db 248 CAATCTGACAAACAGATGCAAGGGAGCAAGATCCATCATGTTTAAATGGAGGAACAGCA 189  
 QY 443 LeuAlaLysArgMetTyrAsnGlnGlnIleLysArgProArgLeuSerProMetGluMet 462  
 Db 188 TTGACGAAGAGATTGTATACCAACAGATGAACAGTCCAAAGTGTTCAGACATGGAATG 129  
 QY 463 ProProValThrSerSerSerTyrSerProIleTyrArgAspArgSerPheProSerGln 482  
 Db 128 CCACACAGACTTCCTCATCTTATCTTCCTACCTACCTCCGCAAGCTTCCCTAGTCAC 69  
 QY 483 ArgAspAspGlnAspGluIleSerAlaLeuValSerSerTyrLeuGlyProSerThr 502  
 Db 68 AGAGAC-----GATGAATATCAGCTTGTGTCAGTAGTTACCTCGAGCCTTCACCA 18  
 QY 503 SerPheProHisArg 507  
 Db 17 GGTTCCTCATCGG 3

RESULT 6 CK255208 931 bp mRNA linear EST 12-DEC-2003  
 LOCUS EST738845 potato callus cDNA library, normalized and full-length  
 DEFINITION Solanum tuberosum cDNA clone POCU48 5' end, mRNA sequence.

ACCESSION CK255208  
 VERSION CK255208.1 GI:39811864  
 KEYWORDS EST.

SOURCE Solanum tuberosum (potato)  
 ORGANISM Solanum tuberosum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.

AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 TITLE Generation of ESTs from potato callus tissue  
 JOURNAL Unpublished (2003)  
 COMMENT Other\_ESTs: EST738846

CONTACT: Robin Buell  
 The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES Location/Qualifiers

1..931  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"  
 /clone="POCCJ48"  
 /tissue\_type="callus"  
 /lab\_host="DH10B-Tona"  
 /clone\_lib="potato callus cDNA library, normalized and full-length"  
 /note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

## ORIGIN

Alignment Scores: 9.64e-43 Length: 931  
 Pred. No.: 580.00 Matches: 125  
 Score: 59.56% Conservative: 65  
 Percent Similarity: 39.18% Mismatches: 106  
 Best Local Similarity: 18.69% Indels: 24  
 Query Match: 14 Gaps: 4  
 DB: CK252381

US-09-890-475-1 (1-609) x CK255208 (1-931)  
 QY 113 LeuSerProProArgAsnValSerValGluThrThrValThrValSerGlnProSer 132  
 Db 5 CTCACCGCCACCGCGCGCAG-ACGGCTGATACACCAC-----CCTTCA 48  
 QY 133 GlnGluIleValProGluThrSerAsnLysProGluGlyArg----- 147  
 Db 49 CGCGCGTCACACCGCTCAGCCTCGATCAACCTCCGAGAAAGAGACTCAATCT 108  
 QY 148 -----MetCysGluLeuMetCysSerLysGlyLeuArgLysTyr 160  
 Db 109 CGGATTTCACAGCTCGAAACCTCTGCAAGACGATGCGCGCAAGGACTTCGAAGTTAC 168  
 QY 161 IleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAlaLeuLys 180  
 Db 169 ATAGTATCGAACTCCAGAGAAACACACTCGCGAGAAAGAACTCCCAAGGCATTGGAA 228  
 QY 181 LeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGly 200  
 Db 229 CTTGCACCCATCGCAGCAAGCTTGTTACTTAGCTGTATGGGAGACTTCTTTGCGAAAGG 288  
 QY 201 ArgArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLe 220  
 Db 289 GGCAAAGCTTTTGTATAGATGTCACAGATGATTGCCAAGCGGAGCATCTGCAATGGTT 348  
 QY 221 LeuGluSerPheLeuLeuMetProAspArgGlyLysValLysIleGluSerTyr 240  
 Db 349 TTGGAGTGTCTTGTGTGATG-----GGTTTCGATGTAATGTAGAGGG 393  
 QY 241 IleLysAspGluAlaGluThrAlaAlaValAlaTyrArgLysArgLeuMetThrGluGly 260  
 Db 394 GTAAAGGAAGAAGCTGCACAAGCAGCAGTAATATGGAGAGAGAGTTTGTGTGAGCAAGA 453  
 QY 261 GlyLeuAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPhe 280  
 Db 454 GGTATAAGAAAGGCTAGTACATGATGCGCGGGATGCTTATCTTATGGGTGTTTT 513  
 QY 281 GlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySer 300  
 Db 514 GGGATTCCACAATTTTTTACAAAGCAGCAGCTTAGGAGCTTGAATTCGAGTCAGTAATATC 573  
 QY 301 AsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyTyr 320  
 Db 574 AGGGAGATTTCTACTGCACTCAGGAGATCAAAATGTCTAATGGAAAGATTCAGAAATA 633  
 QY 321 ValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyrThrPhe 340  
 Db 634 ATAGAGGGCATGGTGAACATAAGATGGAAGCTGCTGTAGATGTTGCCCTATACTTTT 693  
 QY 341 GlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLys 360  
 Db 694 GGGTTTGGAGAAAAATGTAGCCCTTTCGAAATTTTGTCACTATTTTGAAGAACTTAAA 753  
 QY 361 GluSerPheGluArgAlaLysArgLysAlaGln--SerProLeuAlaPheLysGluAla 379  
 Db 754 GAATCCTGAGAAAAAGAAATGCCATCGCATGGTTCACATGCTGTGTGAATGAAGCA 813  
 QY 380 AlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeu 399  
 Db 814 AACAGAGAGGAATTTCTACTATAATCTGCTATTGAATCTTTGGAAGCCCAACCATTT 873  
 QY 400 AspProAlaLysGluLeuProGlyTyrGlnIleLysGlnIleValSerLeuGlu 418  
 Db 874 GATCCCTCAAGCTTATTTCACGATTCGGTATCAGTACAGAAATATAGCTTGGAG 930

## RESULT 7

CK252381

LOCUS

DEFINITION

ACCESSION

VERSION

GI:39806317

CK252381

943 bp mRNA linear EST 12-DEC-2003

ES736018 potato callus cDNA library, normalized and full-length

Solanum tuberosum cDNA clone POCBU72 5' end, mRNA sequence.



tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (*Pseudomonas syringae* pv. tomat. 12 hr; *Xanthomonas campestris* pv. campestris 12 hr, 18hr; *Pseudomonas syringae* pv. phaseolicola 18hr, and *Xanthomonas campestris* pv. vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts.<sup>24</sup>

## ORIGIN

Alignment Scores:	
Pred. No.:	5,398-40
Score:	550.50
Percent Similarity:	61.41%
Best Local Similarity:	40.60%
Query Match:	17.74%
DB:	14
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:

US-09-890-475-1 (1-609) x CK286517 (1-962)

**Qy**      126 valThrValSerGlnProSerGlnGluIleValProGluThrSerAsnLys-ProGluGI 145  
          :::| | | | | : : : : :  
**Db**      87 ATCACTACCGGGAGCCGCCGAAGAAAAAGACTCAAAGAACCACCTTCCAATTCACAGCT 146

**Qy** 145 yGlyArgMetCysGluLeuMetCysSerLysGlyLeuAArgLysTrIleTyrAlaAsnII 165  
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::  
**Db** 147 CGAAACCTCTGCAGACAGATGTCCGGCAAGGGACTTCGAACTTATATAATATCGCAACT 206

Qy 165 e-----SerAspGlnAlaLysLeuMetGlucIuIlleProSerAlaIeuLysLeuAl 182  
:  
:: :||| :||| :||| :||| :||| :||| :||| :|||  
D**b** 207 GCACACTCCAGAGCAAGAAACAACACTGGCGGAACAACCTCCCAGAGGTATTAAGCTTGC 266

Qy 182 alysgluProAlaIalysPheValIeuAspCysIleGlyLysPheTrLeuGlnGlyArgAr 202  
Db 267 ACCTAACGACGACGAAGCTTGTACTACAGCTCGCTGGGAGACTTTTACGCTAAAGGGGCAA 326

**Qy** 202 gAlaPheThrLysGluSerProMetSerAlaArgGlnValSerLeuIleuLeuG1 222  
:  
|||::: ||| |||::: ||| |||::: |||  
**Dd** 327 AGTCTTAATAGACACCGCAGATGATTTCCTATAGGGAGCATCCGCATTGGTTTGA 386

Qy 222 uSerPheIeuLeuMetProAspArgGlyLysGlyLysValLysIleGluSerTrpIleLy 242  
 Db 387 GTCCTTTTATTGANG-----GGTTTGACGAAATGATGATCGTTTC 431

[illegible]

**Qy**            262 uAlaAAlaAGluIuysMetAspAlaArgGlyLeuLeuValAlaCysPheGlyVa 282  
:  
**Dd**            492 AAAAAAGGCAAATAAAAATCGATGCCGGGGTTTGTCTTCCTTAATGGTGTTTTGGGAT 551

[illegible]

Qy 302 uIleAlaGlyAlaIeuLysArgSerGlnPheLeuValProMetValSerGlyIleValG1 322  
Db 612 GATTAAAGTGCTCTTAGAGAGTCAAAATGTGCTCATCGAAAAGATCCAGAAATATAGA 671

Qy 322 uSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetVal.TyrThrPheGlyMe 342  
Db 672 GGGCATGTTGAAGCATAAAGATGGAGGCTTGTGATGTTGCTTATACTTTTGGCGT 731

Qy 342 tGIuAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluse 362  
 Db 732 TGAGGAAAAGTTAGCCCTTACACAAATTTGTGCATCGTTTTTACGAAATTTAAGAATC 791

**Qy**      362 r p h e g l u a t a l a t y s a r g l y s a l a g l n --- s e r p r o l e u a l a p h e l y s g l u a l a t h 381  
          | : : : : | | | | : : : |  
**D b**      792 G C T G A A G A N A A G G A A A T G G A A A T C A C A T G G T T C A C A T G T T G T T C A A T A A G C A A C A A 851

Qy	381	lRyGlnLeuAlaValLeuSerValMetGlnCysMetGluThrHisLysLeuAspPr	401
		::: :::	
Db	852	GAGGGAAATGTCTACTATGAAATCTGTCACTGAATGTTTGAGGCC-CATTAACGTTGATCC	910
		::: :::	
Qy	401	oAlaLysGluLeuProGlyTrpGlnIleLysGluGlnIleValSerLeuGlu	418
		::: :::	
Db	911	CCCTCAAGCTTATCCACGGTTTCGTATCACTGCACAGAGTTTGAGAGCTTGAG	962
		::: :::	

## RESULT 9

B77633/C	B77833	635 bp	DNA	linear	GSS 16-JAN-1998
LOCUS	T29E23	TR	Arabidopsis thaliana	genomic clone	T29E23, genomic
DEFINITION					survey sequence.

ACCESSION  
VERSION

KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
GSS.

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosoids I; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 635)

REFERENCE

**AUTHORS**  
Kounley, S.D., Field, C.E., Bass, S., Limer, K., Limer, A.,  
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Adams, M.D. and  
Venter, J.C.

111115  
 A BAC end sequence database for identifying minimal overlaps in  
 Arabidopsis genomic sequencing. Update 3  
 JOURNAL  
 Unpublished (1997)  
 COMMENT  
 Other GSSs: T39E23TF

Contact: Steve Kounis  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA

Email: [roundsley@tiger.org](mailto:roundsley@tiger.org)  
 Seq primer: M13 Reverse  
 Tel: 301 838 0200  
 Fax: 301 838 0208

```

Ciabs: SAC ends
High quality sequence stop: 635.
Location/Qualifiers
1. .635
FEATURES
SOURCE
```

```

/organism="Haemophilus influenzae"
/mol_type="Genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"

```

```

/crunch
/size=1000000
/sex="hermaphrodite"
/clone_lib="TAMU"
/notes="Vector: BeloSACII; Site 1: HindIII; Site 2:

```

## ORIGIN

Alignment Scores: 1.34e-39 Length: 635  
Pred. No.: 543.00 Matches: 98  
Score:

```

Percent Similarity: 100.00%
Percent Local Similarity: 100.00%
Mismatch: 0
Mismatches: 0
Indels: 0
Gaps: 0
DB: 28
Query Match: 17.49%

```

US-09-890-475-1 (1-609) x B77833 (1-635)  
Qy 512 ProGluTyrMetValProLeuproHisGlyGlyLeuGlyArgSerValTyrAlaTyrclu S31

Db	CCGGAATATATGGTTCCACTTCACATGGTGGTTAGGAAGAAGTATATGCATATGAA	570
Qy	HisLeuAlaProAsnSerTyrSerProGlyHisGlyHisArgLeuHisArgGlnTyrSer	551

Db	559	CACTGGCCCCAAATTCATATTTCCAGGTACGGACATAGACTTCATCGACAGTACTCT	S10
Qy	552	ProSerLeuValHisGlyGlnArgHisProLeuGlnTyrSerProIleHisGlyGln	S71

Db  
509 CCGTCTTTGGTTTCACGGACAGACATCCATACAGTACTCTCTCCATTCATGGACAA 450

```

QY 572 GlnGlnLeuProTyrGlyLeuValTyrArgHisSerProSerGluGluArgTyr 591
Db 449 CAACAGTATACCATATGTTATCAAAAGGTTTACAGACATTCACCATCTGAAGAAAGATAT 390
QY 592 LeuGlyLeuSerAsnGlnArgSerProArgSerAsnSerSerLeuAspProLys 609
Db 389 TTGGGTTTATCAATCAAAAGGTTCTCTCGAGTAACATCATATTAGACCCCAAA 336

RESULT 10
CK243005 905 bp mRNA linear EST 12-DEC-2003
LOCUS Solanum tuberosum cDNA clone F0CA208 5' end, mRNA sequence.
DEFINITION CK243005
ACCESSION CK243005
VERSION CK243005.1 GI:39787192
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 905)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Other ESTs: EST726641 EST726643 EST726644
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from TIGR via potato@igr.org
Seq primer: ANT TAG GTG ACA CTA TAG.
FEATURES
source
location/Qualifiers
1..905
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCA208"
/tissue_type="callus"
/lab_host="DH10B-TconA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."
ORIGIN
Alignment Scores:
Pred. No.: 1.18e-38 Length: 905
Score: 535.50 Matches: 118
Percent Similarity: 57.38% Conservative: 57
Best Local Similarity: 38.69% Mismatches: 107
Query Match: 17.25% Indels: 23
DB: 14 Gaps: 4

US-09-890-475-1 (1-609) x CK243005 (1-905)
QY 109 HisGlnProMetLeuSerProProArgAsnValSerValGluThrVal 128
Db 1 CATCTCCGATGAGAAATCTGGCTGCTCCGATATCCAGACTCCCAACCCACCGCG 60
QY 129 SerGln-----ProSerGlnGluLeuValProGluThrSerAsnLysPro 143
Db 61 CCGCAGACGGCTGATACCAACCCCTTCACCGCGCTCACCAACCGTTCGAAATCAAAC 120
QY 144 GluGlyGlyArg-----MetCysGluLeu 151
Db 121 TCATCCGAGAAAAGAACTCAAAATCTCGGATTCACAGCTCGAAACCTCTGCAAGCG 180
QY 152 MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeu 171

```

```

Db 181 ATGTGCGGCAAGGAGCTTCGAAGTTACATAGTATCGAAATCCCGAGAAAAACACACTG 240
QY 172 MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp 191
Db 241 CGCGAAGAACTCACCAAGGCAATTCGAATTCACCAATTCGAGCAAAAGCTTGACTTAGC 300
QY 192 CysIleGlyLysPheThrLeuGlnGlyArgArgAlaPheThrLysGluSerProMetSer 211
Db 301 TGTATGGGAGACTCTTCGGAAGGGGCAAGCTTTTGATAGAAATGCACAGATGATT 360
QY 212 SerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGly 231
Db 361 GCCACAAGGAAGCATCTGCATTTGTTGGAGTGTCTTCTTGTGATG----- 408
QY 232 LysGlyLysValLysIleLeuSerTrpIleLysAspGluAlaGluThrAlaAlaValAla 251
Db 409 ---GGTTTCGATGTAATTGATGAAGGGTAAAGGAAGAGCTGCACAGCAGCAATATA 465
QY 252 TrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaAlaLysMetAspAlaArg 271
Db 466 TGGAGAAAGAGTTGTTGATGACGAGGTATAAGAAAGGCTAGTACAATGATGCGCG 525
QY 272 GlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeu 291
Db 526 GGAATGCTATTGCTATTGCTGTTTGGGATTCACAAATTTTTCACAAATGAGGACGTT 585
QY 292 LeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGln 311
Db 586 AGGCACTTGATTCGAGTGAATATACAGGAGATTTCTACTGCATCAGGAGATCAAAAT 645
QY 312 PheLeuValProMetValSerGlyIleValGluSerSerIleLysArgGlyMetHisIle 331
Db 646 GTGCTAATGCAAAAGATTCAGAAATATATAGAGGCGCATGCTGAAGCATAGATGGAAGTT 705
QY 332 GluAlaLeuGluMetValTyrThrPheGlyMetGluAspLysPheSerAlaAlaLeuVal 351
Db 706 GATGCTGATAGTGTGCTTATCTTTGAGTTGAGAAATATGAGCCCTTCGCAAAAT 765
QY 352 LeuThrSerPheLeuLysMetSerLysGluSerPheGluArgAlaLysArgLysAlaGln 371
Db 766 TTGTCATCATTTTATAGAAACTTAAAGAAATCGCTGAGAAAAAGAAATGCCAATCGCAT 825
QY 372 ---SerProLeuAlaPheLysGluAlaAlaThrLysGlnLeuAlaValLeuSerSerVal 390
Db 826 GGTTCATAGCTGCTGTTGTAATGAAGCAACAGAGGGAATTTGCTACTATGAATCTGTC 885
QY 391 MetGlnCysMetGlu 395
Db 886 ATTCGAATGTTGGAA 900

RESULT 11
CK294055 913 bp mRNA linear EST 15-DEC-2003
LOCUS Eukaryota; Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION EST756769 Nicotiana benthamiana cDNA clone NBMC57 5',
normalised, full-length Nicotiana benthamiana cDNA clone NBMC57 5',
end, mRNA sequence.
ACCESSION CK294055
VERSION CK294055.1 GI:39877075
KEYWORDS EST.
SOURCE Nicotiana benthamiana
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 913)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org

```



Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GTG ACA CTA TAG.

# FEATURES

source  
1. 913  
Location/Qualifiers  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMCK57"  
/tissue\_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

## ORIGIN

Alignment Scores:  
Pred. No.: 5,928-38 Length: 913  
Score: 528.00 Matches: 116  
Percent Similarity: 60.50% Conservative: 54  
Best Local Similarity: 41.28% Mismatches: 101  
Query Match: 17.01% Indels: 11  
DB: 14 Gaps: 3

US-09-890-475-1 (1-609) x CK294055 (1-913)

Qy 126 ValThrValSerGlnProSerGlnGluLeuValProGluThrSerAsnLys-ProGluG1 145  
Db 87 ATCACTACCGGAGCGCGGAGAGAAAGAAAGACTCAAAAGAAACCCATCCAAATCCACAGC 146  
Qy 145 yGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsn11 165  
Db 147 CGAAACCCCTCTCGACAGCATGTGGCGGAGGAGCTTCGACTTATATATATATATCCACT 206  
Qy 165 e-----SerAspGlnAlaLysLeuMetGluGluLeuProSerAlaLeuLysLeuAl 182  
Db 207 GCCACTCCAGAGCAAGAAACACACACTGGCGAACACACTCCAGAGGTATTAAAGCTTGC 266  
Qy 182 alysgluProAlaLysPheValLeuAspCysileGlyLysPheTyrLeuGlnGlyArgAr 202  
Db 267 ACCTAACGCGAGCAAGCTTGTACTCAGCTGCGTGGGAGCTTTTACGCTAAAGGGGCAA 326  
Qy 202 gAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLeuG1 222  
Db 327 AGATCTTATATAGGACACCGCAGATGATTTCTTATAGGGAAGCATCCGCAATGGTTTGA 386  
Qy 222 uSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysileGluSerTrpIleiy 242  
Db 387 GTGTCTTTTATTGATG-----GGTTGACGAAGAAATGATGAAGTGGTTTC 431  
Qy 242 sAspGluAlaGluThrAlaAlaValAlaTrpArgLysArgLeuMetThrGluGlyGlyLe 262  
Db 432 AAAGAGAGTGAACAGCGCGGTATATATGAGAGAGAGATGGTTGTATGAAGAGGTAT 491  
Qy 262 uAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuValAlaCysPheGlyVa 282  
Db 492 AAAAAAGGCAATAAATGATGCGCGGGGTTTGTGTCTTATTGGGTGTTTGGGAT 551  
Qy 282 lProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnG1 302  
Db 552 TCACAAAGCTTTTGAAGATGAGGATGTTAGGAGTTTGTTCGGAAGAGTAAACATCAGGG 611  
Qy 302 uileAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValG1 322

Db 612 GATTAAGAGTGTCTTAGGAGATCAAAATGCTCATCGAAAGATTCGAGAATATATAGA 671  
Qy 322 uSerSerIleLysArgGlyMetHieIleGluAlaLeuGluMetValTyrThrPheGlyMe 342  
Db 672 GGGCATGGTGAAGCATAAGATGAAGTTGAGGCTGTGATGTTGCTTATCTATCTTTGGCGT 731  
Qy 342 tGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSe 362  
Db 732 TGAGGAAGAAAGTAGCCCTTACAAATTTTGTCATCGTTTACATGAATTTAAAGAATC 791  
Qy 362 rPheGluArgAlaLysArgLysAlaGlnSerProLeuAlaPheLysGluAlaAlaThrIy 382  
Db 792 GCTGAAGAAAGAAAGAAATGGAATCACATGCTTCACATGATGTTGT-GAATAAGCAACAA 850  
Qy 382 s---GlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspPr 401  
Db 851 GAGGGAATTTGCTACTACTGAATCTGTCAATGTTTGGAGGCCCATACGTTGATTC 910  
Qy 401 o 401  
Db 911 C 911

## RESULT 12

CK288931

## LOCUS

DEFINITION

EST751653 Nicotiana benthamiana mixed tissue cDNA library,

normalized, full-length Nicotiana benthamiana cDNA clone NBMK18 5'

end, mRNA sequence.

CK288931

VERSION

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Nicotiana.

1 (bases 1 to 945)

Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,

Staskawicz,B., Jin,H. and Baker,B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST751654

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

1..945

/organism="Nicotiana benthamiana"

/mol\_type="mRNA"

/db\_xref="taxon:4100"

/clone="NBMK18"

/tissue\_type="abiotic and biotic stress-treated leaves,

callus tissue and root tissue"

/lab\_host="DH10B-Tona"

/clone\_lib="Nicotiana benthamiana mixed tissue cDNA

library, normalized, full-length"

/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Nicotiana benthamiana

tissues that include callus, roots from liquid culture

grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),

cold-stressed leaves (5 C 3 hr, 6hr), and pathogen

challenged leaves (Pseudomonas syringae pv tomato 12 hr;

Xanthomonas campestris pv campestris 12 hr, 18hr;

Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas

campestris pv vesicatoria 18hr). RNA was isolated from

these tissues and pooled in approximately equal molar

amounts."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.07e-37 Length: 945  
 Score: 525.50 Matches: 116  
 Percent Similarity: 61.13% Conservative: 57  
 Best Local Similarity: 40.93% Mismatches: 100  
 Query Match: 16.93% Indels: 11  
 DB: 14 Gaps: 3

US-09-890-475-1 (1-609) x CK288931 (1-945)

QY 126 ValThrValSerGlnProSerGlnGluLeuValProGluThrSerAsnLys-ProGluG1 145  
 DB 108 ATCACTACCGGCGCGCGAAGAAAGACTCAAGAAACCACTTCCAAATCCACAGCT 167  
 QY 145 yGlyArgMetCysGluLeuMetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnI1 165  
 DB 168 CGAAACCCCTCTCAAGACGATGTCGGCAAGGATCTCGAACTTATATAATATCGCAACT 227  
 QY 165 e-----SerAspGlnAlaLysLeuMetGluGluLeuProSerAlaLeuLysLeuAl 182  
 DB 228 GCACCTCCAGAGCAAGAAACACACTGCGCAACCACTCCAGAGGTATTAAAGCTTGC 287  
 QY 182 aLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeuGlnGlyArg 202  
 DB 288 ACTAACCGCAGCGAAGCTTGTACTCAGCTGCGTGGGAGACTTTTACGCTAAAGGGGCAA 347  
 QY 202 galapheThrLysGluSerProMetSerSerAlaArgGlnValSerLeuLeuLysLeuG1 222  
 DB 348 AGATCTTAATAGGACACGACATGATTCCTATAGGAGAGATGGTGTGATGAAGAGTAT 407  
 QY 222 uSerPheLeuLeuMetProAspArgGlyLysValLysIleGluSerTyrIleLys 242  
 DB 408 GTGCTCTTTATTGATG-----GGTTTGAACGAAATGATGAAGTGGTTTC 452  
 QY 242 shpGluAlaGluThrAlaAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 262  
 DB 453 AAAAGAAAGTGAACCAAGCGCGGTATATGGAGAGAGATGGTGTGATGAAGAGTAT 512  
 QY 262 uAlaAlaAlaGluLysMetAspAlaArgGlyLeuLeuLeuLeuValAlaAlaCysPheGlyVa 282  
 DB 513 AAAAAGGCAATAAATGATGCGCGGGTGTGCTGTGCTTATTTGGTGTGTTGGGAT 572  
 QY 282 lProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSerGlySerAsnG1 302  
 DB 573 TCCACAAGCTTTTGAAGATGAGATGTTAGGAGATTTGATTCGAGAGAGTAACATCAGGGG 632  
 QY 302 uIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSerGlyIleValG1 322  
 DB 633 GATTAAAGTCTCTTAGGAGATCAATGTGCTCATCGAAAGATTCAGAAATAATAGA 692  
 QY 322 uSerSerIleLysArgGlyMetHisIleGluAlaLeuMetValTyrThrPheGlyWe 342  
 DB 693 GGGCATGTTGAAGCATAAGATGGAAGTGTGAGCTGTGATGTTGCTTATCTTTGGCGT 752  
 QY 342 tGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMetSerLysGluSe 362  
 DB 753 TGAGGAAAAGTTAGCCCTTACCAATTTGTGATCTGTTTACATGAATTTAAAGATC 812  
 QY 362 rPheGluArgAlaLysArgLysAlaGln---SerProLeuAlaPheLysGluAlaAlaTh 381  
 DB 813 GCTGAAGAAAAGGAAATGGAATACATGATGTTTGTGTAATAAAGCAACAA 872  
 QY 381 rLysGlnLeuAlaValLeuSerSerValMetGlnCysMetGluThrHisLysLeuAspPr 401  
 DB 873 GAGGGAAATGCTACTATGAATCTGTCTCAATGTTGGAGCC-CATAACGTTGATCC 931  
 QY 401 cAlaLys 403  
 DB 932 CTCAAG 938

RESULT 13  
 CD836567 684 bp mRNA linear EST 10-JUL-2003  
 LOCUS

DEFINITION BN45.049C12F020103 BN45 Brassica napus cDNA clone BN45049C12, mRNA  
 sequence.  
 ACCESSION CD836567  
 VERSION CD836567.1 GI:32518507  
 KEYWORDS EST.  
 SOURCE Brassica napus (rape)  
 ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 REFERENCE 1 (bases 1 to 684)  
 AUTHORS Genoplante,  
 TITLE Genoplante, a major partnership french program in plant genomics  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Genoplante  
 Genoplante  
 92, rue Henri Rochefort 91025 EVRY CEDEX France  
 Tel: 33 1 69 47 54 00  
 Fax: 33 1 69 47 54 10  
 This sequence has been generated in the framework of the french  
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>  
 and <http://genoplante-info.infobiogen.fr>).  
 FEATURES  
 Location/Qualifiers  
 1..684  
 /organism="Brassica napus"  
 /mol\_type="mRNA"  
 /cultivar="Jet Neuf"  
 /db\_xref="taxon:3708"  
 /clone="BN45049C12"  
 /tissue\_type="seed"  
 /clone\_lib="BN45"

Alignment Scores:  
 Pred. No.: 9.58e-38 Length: 684  
 Score: 523.50 Matches: 110  
 Percent Similarity: 66.98% Conservative: 32  
 Best Local Similarity: 51.89% Mismatches: 49  
 Query Match: 16.87% Indels: 21  
 DB: 14 Gaps: 3

US-09-890-475-1 (1-609) x CD836567 (1-684)

QY 16 AlaAsnProLeuLeuGlnArgHisGlnSerGluGlnArgArgGluLeuProLysIle 35  
 DB 96 TCATCGCGGACCATACACCG-----116  
 QY 36 ValGluThrGluSerThrSerMetAspIleThrIleGlyGlnSerLysGlnProGlnPhe 55  
 DB 117 ---GGAACCGGGCTACAAACAGCGAATCAGCATCGAACATCTAACCATCTCAATTT 173  
 QY 56 LeuLysSerIleAspGluLeuAlaAlaPheSerValAlaValGluThrPheLysArgGln 75  
 DB 174 TTGAATTCGATCGACGATTTAACTGCGTTTTCAGTCGATGACGCTTCAACCGCCAC 233  
 QY 76 PheAspAspLeuGlnLysHisIleGluSerIleGluAsnAlaIleAspSerLysLeuGlu 95  
 DB 234 TACGACGACTTGAACAGCTCTCATGATTACATCAAGAACGCCATTGACTCAGCTCTCAAG 293  
 QY 96 SerAsnGlyValValLeuAlaAlaArgAsnAsnAsnPheHisGlnProMetLeuSerPro 115  
 DB 294 AGCAAGAGCATACCGCGGAGTCTCCTCTCCGATCGCAGTCT-----338  
 QY 116 ProArgAsnValSerValGluThrThrValThrValSerGlnProSerGlnGluIle 135  
 DB 339 CCACGAAACGATGCTTCGGGAGAAACGGTTGCTGCCACAAATCGCCGCAAGAGAGACT 398  
 QY 136 ValProGluThrSerAsnLysProGluGlyGlyArgMetCysGluLeuMetCysSerLys 155  
 DB 399 TGTGAGACAGTAGCGGAGAGGTGGAG-----CGATTGCGGAGTTGATGTGCAGCAA 452  
 QY 156 GlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIle 175

Db 453 GGCTCGGTAGATACATGTACTCGAATATCTCTGACCGAGCTAAGCTGATGTAAGAGCTT 512

Qy 176 ProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLys 195

Db 513 CTTGCGAGCTCTGAAGCTAGCCAGGAGCTGTAACTCTGTGTGGGAATGCAITGGCAAG 572

Qy 196 PheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGln 215

Db 573 TTTTACTTACAGGGCGCAAGCTTATGCGAGTGATTTCCATATGATTCCTCGGAGCGAG 632

Qy 216 ValSerLeuLeuLeuGluSerPheLeuLeuMet 227

Db 633 GTTTCGCTTCTGATCCTGGAGTCTTATCTTCTTAATG 668

RESULT 14

CK249078 883 bp mRNA linear EST 12-DEC-2003

LOCUS EST732715 potato callus cDNA library, normalized and full-length

DEFINITION Solanum tuberosum cDNA clone POCS576 5' end, mRNA sequence.

ACCESSION CK249078.1 GI:39799840

VERSION EST.

KEYWORDS Solanum tuberosum (potato)

SOURCE Solanum tuberosum

ORGANISM

REFERENCE

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

JOURNAL asterids; lamids; Solanales; Solanaceae; Solanum.

COMMENT 1 (bases 1 to 883)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from potato callus tissue

Unpublished (2003)

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

Location/Qualifiers

1..883

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POCS576"

/tissue\_type="callus"

/lab\_host="DH10B-TonA"

/clone\_lib="potato callus cDNA library, normalized and full-length"

/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

ORIGIN

Alignment Scores:

Pred. No.: 3,42e-37 Length: 883

Score: 519.50 Matches: 115

Percent Similarity: 57.91% Conservative: 57

Best Local Similarity: 38.72% Mismatches: 102

Query Match: 16.74% Indels: 24

DB: 14 Gaps: 4

US-09-890-475-1 (1-609) x CK249078 (1-883)

Qy 111 ProMetLeuSerProProArgAsnValSerValGluThrThrValThrValSerGln 130

Db 24 CCCAGACTCCACCCGCCACCGCGCGAG-ACGGCTGATACCACC- 67

Qy 131 ProSerGlnGluLeuValProGluThrSerAsnLysProGluGlyGlyArg- 147

Db 68 CTTTCACCGGGGTACACCGCTCAGCTCGATCAAACTCATCCGAGAAAGACTCA 127

Qy 148 -----MetCysGluLeuMetCysSerLysGlyLeuArg 158

Db 128 AAATCTCGGATTACAGCTCGAAACCCCTCTGCAAGACGATGTGCGCAAGGGACTTCGA 187

Qy 159 LysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeuMetGluGluIleProSerAla 178

Db 188 AGTTACATAGTATCGAAACTCCAGAGAAACACACTGCGCGAAGAACTCACAAGGCA 247

Qy 179 LeuLysLeuAlaLysGluProAlaLysPheValLeuAspCysIleGlyLysPheTyrLeu 198

Db 248 TTGGAACTTGCACCCATGCGACCAAGCTTGTACTTACTGTATGGGAGACTTCTTTGCG 307

Qy 199 GlnGlyArgAlaPheThrLysGluSerProMetSerSerAlaArgGlnValSerLeu 218

Db 308 AAAAGGGGCAAAAGCTTTTGATAAGATGCACAGATGATTCACACAGGAGAGCATCTGCA 367

Qy 219 LeuIleLeuGluSerPheLeuLeuMetProAspArgGlyLysGlyLysValLysIleGlu 238

Db 368 TTGGTTTTGGAGTGTCTTCTTGTGATG-----GGTTTCGATGTATTTGAT 412

Qy 239 SerTTrpIleLysAspGluAlaGluThrAlaAlaValAlaLysArgLysArgLeuMetThr 258

Db 413 GAAGGGGTAAAGGAAGCTGCACAGCAGCAGTAAATATGAGAGAAGGTTTGTTCAT 472

Qy 259 GluGlyGlyLeuAlaAlaAlaGluLysMetAspAlaArgGlyLysLeuLeuValAla 278

Db 473 GAACGAGTATAGAAAGGCTAGTACAATGATGCGCGGATTCCTATTGCTATTATGGG 532

Qy 279 CysPheGlyValProSerAsnPheArgSerThrAspLeuLeuAspLeuIleArgMetSer 298

Db 533 TGTTTTGGATTCACAAATTTTTTACAAATGAGGAGCTTAGGACCTTCTGATTCGAGT 592

Qy 299 GlySerAsnGluIleAlaGlyAlaLeuLysArgSerGlnPheLeuValProMetValSer 318

Db 593 AATATCAGGGAGATTCTACTGCATCAGGAGATCAAAATGTCTAATGGAAAAGATTCCA 652

Qy 319 GlyIleValGluSerSerIleLysArgGlyMetHisIleGluAlaLeuGluMetValTyr 338

Db 653 GAATAATAGAGGCGCATGTGAGCATAGATGGAAGTGTGCTGTAGATGTTCCCTAT 712

Qy 339 ThrPheGlyMetGluAspLysPheSerAlaAlaLeuValLeuThrSerPheLeuLysMet 358

Db 713 ACTTTTGGAGTTGAGGAAAAATGTAGCCCTTGCAGAAATTTTGTCAATCATTTTTAGAGAA 772

Qy 359 SerLysGluSerPheGluArgAlaLysArgLysAlaGln---SerProLeuAlaPheLys 377

Db 773 CTTTAAAGAAATCGCTGAAGAAAAAAGAAATGCCAATGCCATGCTTCACATGCTTGTGAT 832

Qy 378 GluAlaAlaThrLysGlnLeuAlaValLeuSerSerValMetGlnCysMet 394

Db 833 GAAGCAACACAGAGGGAATGTCTACTATGAAATCTGTCAATGAATGTTTG 883

RESULT 15

CK243006

LOCUS

DEFINITION

EST726643 potato callus cDNA library, normalized and full-length

Solanum tuberosum cDNA clone POCA208 5' end, mRNA sequence.

ACCESSION CK243006

VERSION CK243005.1 GI:39787194

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM

REFERENCE

AUTHORS

TITLE Generation of ESTs from potato callus tissue

JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST726641 EST726642 EST726644

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

CK243006 904 bp mRNA linear EST 12-DEC-2003

EST726643 potato callus cDNA library, normalized and full-length

Solanum tuberosum cDNA clone POCA208 5' end, mRNA sequence.

ACCESSION CK243006

VERSION CK243005.1 GI:39787194

KEYWORDS EST.

SOURCE Solanum tuberosum (potato)

ORGANISM

REFERENCE

AUTHORS

TITLE Generation of ESTs from potato callus tissue

JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST726641 EST726642 EST726644

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Search completed: February 28, 2004, 05:48:12  
Job time : 3467 secs

US-09-890-475-1 (1-609) x CK243006 (1-904)

QY	109	HisGlnProMetLeuSerProArgAsnAsnValSerValGluThrThrValThrVal	128
DB	2	AAATCTCCGATGGAGAATCTGGCTGGCTCTCCGATATCCAGACTCCACCGCCACCGCG	61
QY	129	SerGln-----ProSerGlnGluIleValProGluThrSerAsnLysPro	143
DB	62	CCGACAGCGGTGATACCAACCCCTTCACCGCGCTCACACCGTCAGCTCGAATCAAC	121
QY	144	GluGlyGlyArg-----MetCysGluLeu	151
DB	122	TCATCCGAAGAAAAAGACTCAAAATCTCCGATTTCACAGTTCGAAACCCCTCTCGACAGC	181
QY	152	MetCysSerLysGlyLeuArgLysTyrIleTyrAlaAsnIleSerAspGlnAlaLysLeu	171
DB	182	ATGTGCGGCAAGGACTTCGAAGTTTACATGATGTCAAACTCCACAGAAAAACACACTG	241
QY	172	MetGluGluIleProSerAlaLeuLysLeuAlaLysGluProAlaLysPheValLeuAsp	191
DB	242	CGCGAAGAACTCACCAAGCATTTGCAATTTGCCCAATGCGCAAGACTTTGACTTAGC	301
QY	192	CysIleGlyPheTyrLeuGlnGlyArgAlaPheThrLysGluSerProMetSer	211
DB	302	TGTATGGGAGACTCTCTTCGAAAAAGGGCAAGCTTTTGATAAGATGCACAGATGATT	361
QY	212	SerAlaArgGlnValSerLeuLeuIleLeuGluSerPheLeuLeuMetProAspArgGly	231
DB	362	GCCACAAAGGAAGACATCTGCATTTGTTTGAGTGTTCCTGTGTATG-----	409
QY	232	LysGlyLysValLysIleGluSerTrpIleLysAspGluAlaGluThrAlaAlaValAla	251
DB	410	---GGTTTCGATGTAATTCAGGGGTAAAGGAAGAGCTGCACACAGCAGCAGTAATA	466
QY	252	TrpArgLysArgLeuMetThrGluGlyGlyLeuAlaAlaGluLysMetAspAlaArg	271
DB	467	TGAGAAAGAAGTTTCTTGATGAACAGAGTATGAAGAAGGCTAGTACAAATGGATGCGCG	526
QY	272	GlyLeuLeuLeuValAlaCysPheGlyValProSerAsnPheArgSerThrAspLeu	291
DB	527	GGATTCGTATTGGTATTCCGGGTGTTTGGGATTCACAAATTTTTTCAATGAGCAGCT	586
QY	292	LeuAspLeuIleArgMetSerGlySerAsnGluIleAlaGlyAlaLeuLysArgSerGln	311
DB	587	AGGCATTTGATTCGATGAGTGAGTAAATATCAGGGAGATTTCTACTGCCTCAGGAGTCAAT	646
QY	312	PheLeuValProMetSerGlyIleValGluSerSerIleLysArgGlyMetHisIle	331